****

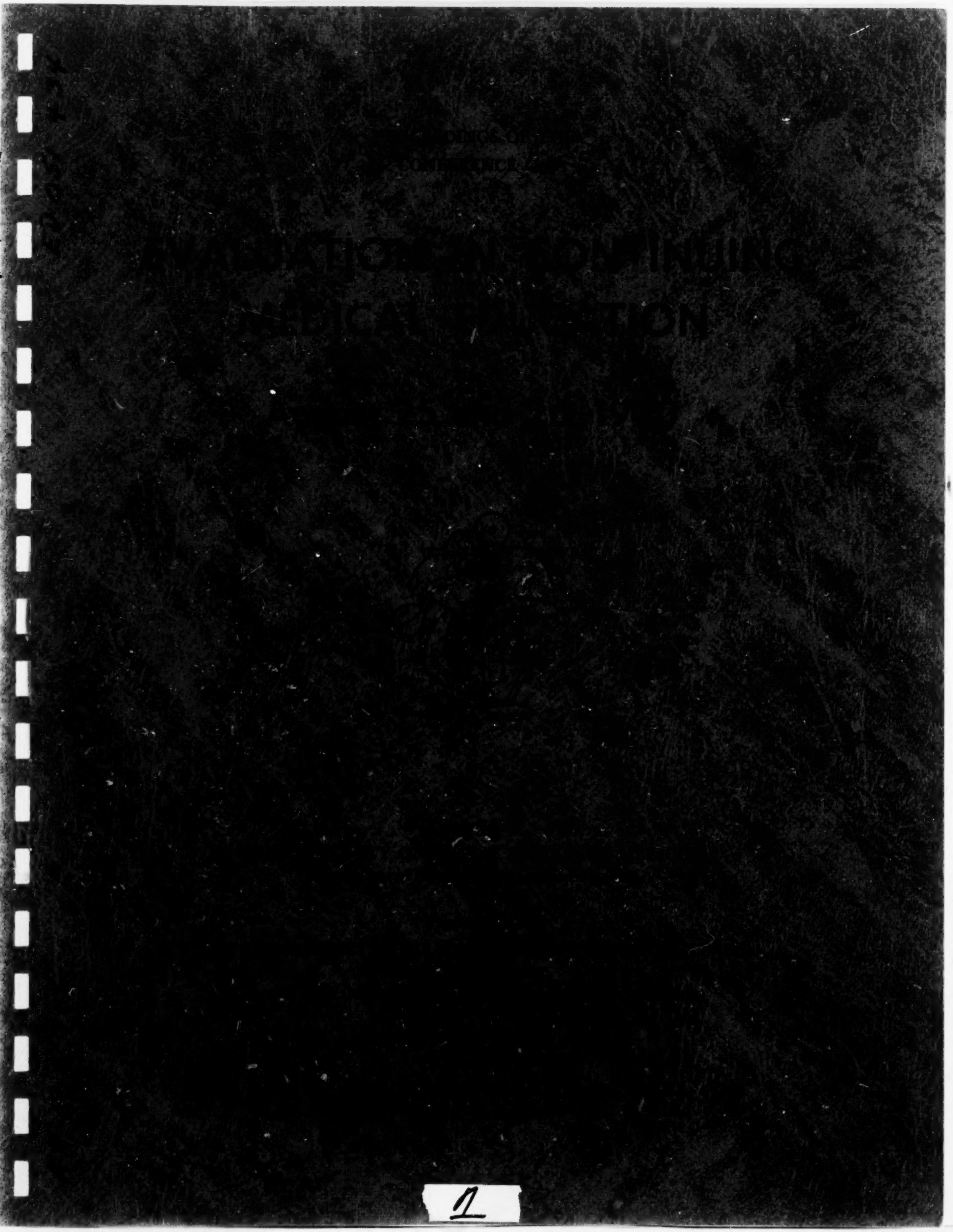
DOCUMENT RESUME

ED 051 434 AA 000 706

AUTHOR	Rising, Jesse D., Ed.
TITLE	Proceedings of the Conference on Evaluation in Continuing Medical Education, August 25-26, 1970.
INSTITUTION	Kansas Univ., Kansas City. Medical Center.
SPONS AGENCY	Public Health Service (DHEW), Washington, D.C. Bureau of Health Professions Education and Manpower Training.
PUB DATE	Aug 70
NOTE	102p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Adult Education, *Conference Reports, *Evaluation, *Health Personnel, *Professional Continuing Education

ABSTRACT

The sole objective of the conference was to assemble experts in the various fields of continuing education for health to discuss the evaluation of their programs. Included were different types of producers of various kinds of courses or programs, adult educationalists, professional evaluators, financers of programs, etc. Particular attention was given to defining the problems encountered in evaluation and to outline possible solutions to them, with the goal of recording these deliberations for the benefit of newcomers to the field. No effort was made to arrive at any resolutions, conclusions, or formal recommendations. Instead, an attempt was made to indicate the complexities of the field of evaluation of continuing medical education efforts and to help those with less experience or with a more restricted point of view. (DAS)



1

PROCEEDINGS OF THE CONFERENCE ON

E V A L U A T I O N

I N

C O N T I N U I N G

M E D I C A L E D U C A T I O N

August 25 and 26, 1970

Edited by

Jesse D. Rising, M.D.

Presented by

THE UNIVERSITY OF KANSAS SCHOOL OF MEDICINE
Department of Postgraduate Medical Education

Supported by

THE DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Bureau of Health Professions Education and Manpower Training
Contract No. NIH 70-4137

Offered at

THE UNIVERSITY OF KANSAS MEDICAL CENTER
Kansas City, Kansas

Table of Contents

	Page No.
Conference Consultants, Consultant Planning Group, Project Officer and Project Director	i
Conference Participants	ii
Preface	ix
FIRST PLENARY SESSION: PROBLEMS AND APPROACHES	1
Educator Attitudes to Evaluation -- Dr. Clement R. Brown	2
Physician Attitudes to Evaluation -- Dr. Neal A. Vanselow	8
Inappropriate or Unrealistic Objectives -- Dr. Stephen Abrahamson	20
Costs of Evaluation -- Mr. William T. Herzog	26
Measurement Problems -- Dr. Hugh M. Petersen	36
GROUP DISCUSSIONS -- Conference participants were assigned by cross-section of backgrounds to group discussions with conference consultants.	
SECOND PLENARY SESSION: GROUP DISCUSSION REPORTS	42
THIRD PLENARY SESSION: REACTION OF CONFERENCE PARTICIPANTS TO GROUP DISCUSSION REPORTS	63
Appendix: Evaluating the Evaluators -- Dr. Ethel Nurge	89

CONFERENCE CONSULTANTS

STEPHEN ABRAHAMSON, Ph.D., Director
Division of Research in Medical Education
University of Southern California School of Medicine
Los Angeles, California

CLEMENT R. BROWN, JR., M.D., Director
Department of Medical Education
Chestnut Hill Hospital
Philadelphia, Pennsylvania

WILLIAM T. HERZOG, M.S.P.H., Director
Continuing Education in Health Sciences
University of North Carolina School of Medicine
Chapel Hill, North Carolina

HUGH M. PETERSEN, Ph.D.
Assistant Professor of Research in Medical Education
University of Southern California School of Medicine
Los Angeles, California

NEAL A. VANSELOW, M.D.
Professor of Postgraduate Medicine
Assistant to the Chairman, Department of Postgraduate Medicine
University of Michigan Medical School
Ann Arbor, Michigan

CONSULTANT PLANNING GROUP

GAIL BANK, Ph.D.
Associate Professor of Extension Education
Director, Continuing Medical Education, Extension Division
University of Missouri-Columbia
Columbia, Missouri

CLEMENT R. BROWN, JR., M.D., Director
Department of Medical Education
Chestnut Hill Hospital
Philadelphia, Pennsylvania

IRWIN H. BROWN, M.D.
Clinical Professor of Continuing Education
Director, Postgraduate Medical Education
University of Oklahoma School of Medicine
Oklahoma City, Oklahoma

LEE D. STAUFFER, M.P.H.
Associate Professor and Dean
Executive Director, Postgraduate Education Activities
University of Minnesota School of Public Health
Minneapolis, Minnesota

PROJECT OFFICER

MARILYN K. HUTCHISON, M.D., Acting Chief
Continuing Education Branch
Division of Health Professions Education and Manpower Training
National Institutes of Health
Bethesda, Maryland

PROJECT DIRECTOR

JESSE D. RISING, M.D., Chairman
Department of Postgraduate Medical Education
University of Kansas School of Medicine
Kansas City, Kansas

CONFERENCE PARTICIPANTS

Stephen Abrahamson, Ph.D., Director
Division of Research in Medical Education
University of Southern California School of Medicine
2025 Zonal Avenue
Los Angeles, California 90033

Gail Bank, Ph.D., Executive Director
Continuing Medical Education
Associate Professor of Extension Education
University of Missouri-Columbia
M-176 University Medical Center
Columbia, Missouri 65201

Elizabeth M. Barnett
Paramedical Education Specialist
Administrative Assistant, Continuing Education
New Mexico Regional Medical Program
920 Stanford Drive, N.E., Building 3-A
Albuquerque, New Mexico 87106

John Barson, Ed.D., Associate Director
Office of Medical Education Research and Development
W-29 Owen Graduate Center
Michigan State University
East Lansing, Michigan 48823

Frank W. Bexfield
Associate for Evaluation and Review
Oklahoma Regional Medical Program
Assistant Professor of Continuing Education
University of Oklahoma School of Medicine
828 Northeast 15th Street
Oklahoma City, Oklahoma 73105

John C. Bish, Ed.D., Assistant Director
Educational Programs
American College of Cardiology
9650 Rockville Pike
Bethesda, Maryland 20014

Clement R. Brown, Jr., M.D., Director
Department of Medical Education
Chestnut Hill Hospital
8835 Germantown Avenue
Philadelphia, Pennsylvania 19118

Irwin H. Brown, M.D.
Clinical Professor of Surgery
Associate Professor of Continuing Education
Director, Postgraduate Medical Education
University of Oklahoma School of Medicine
801 Northeast 13th Street
Oklahoma City, Oklahoma 73104

Margaret M. Brown, Research Associate
Kansas Regional Medical Program
3909 Eaton Street
Kansas City, Kansas 66103

Patricia R. Brown, Ph.D., Assistant Director
Allied Health Professions
Illinois Regional Medical Program
122 South Michigan Avenue, Suite 939
Chicago, Illinois 60603

Thomas C. Brown, Ph.D., Director
Postgraduate Education Section
Department of Community and Family Medicine
University of Utah School of Medicine
50 North Medical Drive
Salt Lake City, Utah 84112

Ronald A. Chez, M.D.
Associate Dean, Academic Affairs
Professor of Obstetrics and Gynecology
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania 15213

James L. Coole, Director
Office of Evaluation
American Academy of Pediatrics
1801 Hinman Avenue
Evanston, Illinois 60204

Paul Cudmore, M.D., Assistant Director
Division of Continuing Medical Education
Dalhousie University Faculty of Medicine
6600 South Street
Halifax, Nova Scotia, Canada

C. Wesley Eisele, M.D., Associate Dean
Postgraduate Medical Education
University of Colorado School of Medicine
4200 East Ninth Avenue
Denver, Colorado 80220

William R. Fifer, M.D., Director
University Division
Northlands Regional Medical Program
University of Minnesota Medical School
Box 193, Health Sciences Center
Minneapolis, Minnesota 55455

Raymond J. Garrity, Evaluation Specialist
Division of Continuing Medical Education
Louisiana State University School of Medicine
1542 Tulane Avenue, Room 231
New Orleans, Louisiana 70112

Constance L. Graves, M.D.
Assistant Professor of Anesthesiology
University of Utah School of Medicine
50 North Medical Drive
Salt Lake City, Utah 84112

Richard H. Hansen, Director
Department of Postgraduate Medical Education
University of Wisconsin Medical School
307 North Charter Street
Madison, Wisconsin 53706

Harold E. Henninger, Coordinator
Education and Training
Georgia Regional Medical Program
938 Peachtree Street
Atlanta, Georgia 30309

William T. Herzog, M.S.P.H., Director
Continuing Education in Health Sciences
University of North Carolina School of Medicine
Miller Hall
Chapel Hill, North Carolina 27514

Gerald E. Hughes, M.D., Director
Department of Educational Affairs
American Academy of Pediatrics
1801 Hinman Avenue
Evanston, Illinois 60204

Keith E. Jenkins, Program Evaluator
Northwestern Ohio Regional Medical Program
1600 Madison Avenue
Toledo, Ohio 43624

Douglas T. Johnston, Director
Slow Scan Television Project
University of Wisconsin Medical School
307 North Charter Street
Madison, Wisconsin 53706

Margaret S. Klapper, M.D.
Professor of Medicine and Associate Dean
Director, Division of Continuing Medical Education
University of Alabama School of Medicine
1919 Seventh Avenue South
Birmingham, Alabama 35233

Donald R. Korst, M.D.
Associate Professor of Medicine
University of Wisconsin Medical School
307 North Charter Street
Madison, Wisconsin 53706

Gaetane M. Larocque, Ph.D.
Associate Program Coordinator
Michigan Association for Regional Medical Program
1111 Michigan Avenue
East Lansing, Michigan 48823

Edwin E. Lauder, M.B.A., Administrative Director
Division of Anesthesiology
University of Utah School of Medicine
50 North Medical Drive
Salt Lake City, Utah 84112

Charles W. Lemke, Director of Evaluation
Wisconsin Regional Medical Program, Inc.
110 East Wisconsin Avenue
Milwaukee, Wisconsin 53202

Frank R. Lemon, M.D., Coordinator
Ohio Valley Regional Medical Program
Associate Dean, Office of Continuing Education
University of Kentucky College of Medicine
Lexington, Kentucky 40506

John W. Lester, M.D., Chairman
Committee on Education
Minnesota Academy of General Practice
Mapleton, Minnesota 56065

Edith Leyasmeyer, Ph.D., Coordinator
Continuing Education
Northlands Regional Medical Program
375 Jackson Street
St. Paul, Minnesota 55033

William E. Matory, M.D., Director
Continuing Medical Education
Howard University College of Medicine
Box 95
Washington, D.C. 20001

Robert R. Moutrie, Director
Department of Continuing Education
University of Nebraska College of Medicine
42nd and Dewey Avenue
Omaha, Nebraska 68105

H. Ormond Murphy, M.D., Assistant Dean
Director, Division of Continuing Medical Education
University of British Columbia Faculty of Medicine
Task Force Building
Vancouver 8, British Columbia, Canada

William P. Nelson III, M.D.
Associate Professor of Medicine
Professor of Postgraduate Medicine
Albany Medical College of Union University
47 New Scotland Avenue
Albany, New York 12208

Robert L. Neth, Executive Director
Department of Postgraduate Medical Education
University of Kansas School of Medicine
Kansas City, Kansas 66103

Henry M. Owen, Evaluator
Arkansas Regional Medical Program
4617 North Mississippi
Little Rock, Arkansas 72207

Nicholas Parlette, M.P.H., Chief
Program of Continuing Education in Public Health,
sponsored by Schools of Public Health of the University
of California at Berkeley, University of California
at Los Angeles, University of Hawaii, Loma Linda
University, and the Western Regional Office,
American Public Health Association,
655 Sutter Street, Suite 408
San Francisco, California 94102

Jack M. Perlman, M.D.
Medical Audit Consultant
Commission on Professional and Hospital Activities
1968 Green Road
Ann Arbor, Michigan 48105

Hugh M. Petersen, Ph.D., Assistant Professor
Research in Medical Education
University of Southern California School of Medicine
2025 Zonal Avenue
Los Angeles, California 90033

Richard D. Petre, Ph.D.
Project Evaluation
Kansas Regional Medical Program
3909 Eaton Street
Kansas City, Kansas 66103

Richard L. Reimer, Assistant Director
Continuing Medical Education
Washington/Alaska Regional Medical Program
530 University District Building
Seattle, Washington 98105

Jerry A. Royer, M.D., Assistant to the Dean
Assistant Professor of Postgraduate Education
Indiana University School of Medicine
1100 West Michigan Street
Indianapolis, Indiana 46202

William L. Schmalgemeier, M.D.
Program Associate
Department of Postgraduate Medicine
University of Michigan Medical School
Towsley Center
Ann Arbor, Michigan 48104

Thelma D. Schneider, Research Analyst
Research and Evaluation Section
Kansas Regional Medical Program
3909 Eaton Street
Kansas City, Kansas 66103

John J. Sharry, D.M.D.
Assistant to the Vice President
Director, Office of Learning Resources
University of Alabama in Birmingham Medical Center
1919 Seventh Avenue South
Birmingham, Alabama 35233

William J. Sheeley, M.D., Assistant Coordinator
Continuing Education
Wisconsin Regional Medical Program
110 East Wisconsin Avenue
Milwaukee, Wisconsin 53202

Marilyn Short, Administrative Associate
Division of Continuing Education
Vanderbilt University School of Medicine
Evaluation Consultant
Tennessee Mid-South Regional Medical Program
1100 Baker Building
110 21st Avenue South
Nashville, Tennessee 37203

Michael Soper, M.D.
Fellow in Continuing Education
Harvard Medical School
Boston, Massachusetts 02115

Lee D. Stauffer, M.P.H.
Associate Professor and Dean
University of Minnesota School of Public Health
1325 Mayo Memorial Building
Minneapolis, Minnesota 55455

Richard C. Surber, B.B.A., Assistant Director
Continuing Medical Education
University of Cincinnati College of Medicine
380 Howell Avenue
Cincinnati, Ohio 45220

David P. Taylor, Coordinator
Research and Evaluation
Susquehanna Valley Regional Medical Program
1104 Fernwood Avenue
P. O. Box 541
Camp Hill, Pennsylvania 17011

Ronald J. Thompson, Assistant Director
Continuing Medical Education
Medical College of Ohio at Toledo
3101 Collingwood Boulevard
Toledo, Ohio 43610

Neal A. Vanselow, M.D.
Professor of Postgraduate Medicine
Assistant to the Chairman
Department of Postgraduate Medicine
University of Michigan Medical School
Ann Arbor, Michigan 48104

Roger J. Warner
Director of Planning and Evaluation
Arkansas Regional Medical Program
500 University Tower Building
12th and University
Little Rock, Arkansas 77205

George W. Welsh, M.D., Director
Continuing Education for Health Sciences
University of Vermont College of Medicines
Given Building
Burlington, Vermont 05401

Jack L. Whenry, Ed.D.
Associate Director of Medical Education
Lovelace Foundation for Medical Education and Research
5200 Gibson Boulevard, Southeast
Albuquerque, New Mexico 87108

Walter J. Zeiter, M.D.
Director of Education
The Cleveland Clinic Educational Foundation
2020 East 93rd Street
Cleveland, Ohio 44106

Thomas F. Zimmerman, Ph.D.
Assistant to the Director
Division of Medical Education
American Medical Association
535 North Dearborn Street
Chicago, Illinois 60610

Sylverster H. Zumbrun, M.Ed., Deputy Director
Center for Continuing Education
Metropolitan Washington Regional Medical Program
5019 Portsmouth Road
Fairfax, Virginia 22030

PREFACE

The conference, Evaluation in Continuing Medical Education, was presented by the Department of Postgraduate Medical Education of the University of Kansas School of Medicine with the support of the Bureau of Health Professions, Education and Manpower Training of the Public Health Service, Department of Health, Education, and Welfare on August 25 and 26, 1970. The sole objective of the conference was to assemble experts in various fields of continuing education for health -- different types of producers of various kinds of courses or programs, adult educationalists, professional evaluators, financers of programs, etc. -- to discuss the evaluation of such programs. We wanted to give particular attention to defining the obstacles or problems encountered in evaluation and outline possible solutions to them with the goal of recording their deliberations for the benefit of others -- particularly newcomers to the field -- who might desire to know what those who had been in the "business" longer had to say.

There is no doubt that most of those in attendance learned something (or at least had their perspectives deepened), but that was not the objective of the conference. It was rather to indicate the complexities of the field of evaluation of continuing medical education efforts and to help those with less experience or with a more restricted point of view. For that reason we made no effort to arrive at any "resolutions," conclusions, or formal recommendations. We have tried instead to record the extemporaneous proceedings with maximum faithfulness. Editorial privilege has been used sparingly, being applied only to eliminate irrelevant side remarks (and some very good jokes) and to try to be sure that the meaning of the spoken word was transmitted faithfully on the rare occasions when oral discourse, although clear in itself, lost its meaning on the printed page.

Logistical problems made it difficult to submit a transcript of each discussant's comments to him for approval, and I hope that all of them will pardon me for taking the liberty of presenting their remarks as they were transcribed by the stenotypist.

The Department of Postgraduate Medical Education has recently had critiques of many of its programs by Ethel Nurge, Ph.D., cultural anthropologist. Unfortunately Dr. Nurge was unable to attend the conference, but was kind enough to read the proceedings and comment on them. Her comments are included in Appendix A.

It is obvious that this conference could not have been carried out successfully without the maximum cooperation of all those who attended, and I want particularly to thank the Consultant Planning Group, the Conference Consultants, and the volunteer Group Discussion Reporters for their roles in planning and producing the conference. It is also a pleasure to thank our extremely capable reporter, Donna L. Dunwoody, Dunwoody Convention Reporting, Shawnee Mission, Kansas, who produced a truly superb transcript of the extemporaneous proceedings -- a transcript that made the job of editing a pleasure.

Jesse D. Rising, M.D.
Kansas City, Kansas
October, 1970

FIRST PLENARY SESSION

EDUCATOR ATTITUDES TO EVALUATION

Clement R. Brown, M.D.

To better serve the medical educators with whom I have consulted during the past two years, I have administered a series of instruments to measure their knowledge, skills and opinions relative to continuing medical education program design. I will report briefly on responses from an opinion survey.

I. Current Educator Attitudes to Evaluation of Continuing Medical Education (CME)

I administered the opinion survey to 157 people directly involved with continuing medical education program production. Most were physicians. Most who were nonphysicians had a background in the behavioral sciences. A few were medical record librarians. Included in the group were members of medical school departments of continuing medical education, regional medical programs, and directors of medical education in community hospitals.

The opinion survey consisted of a series of statements concerning continuing medical education with instructions to the respondents to select those statements with which they were in greatest agreement or disagreement. Statements 2, 16 and 17 are most pertinent to the present discussion. Statement No. 2 -- "Success of CME programs must be measured in terms of improved patient care." One-hundred forty-nine CME program producers responded to this statement -- 135 or 90 percent agreeing, and only 14 or 10 percent disagreeing with the statement.

Statement No. 16 -- "It is possible to measure the effect of CME programs on patient care, and also to determine the need for CME programs by measuring patient care." One-hundred thirty-three responded to this statement -- 118 or 89 percent agreeing, and only 15 or 11 percent disagreeing.

Statement No. 17 -- "The true success of a CME program can best be judged by a "before" and "after" examination (knowledge test) of the participants." One-hundred thirty-five responded to this statement -- only 17 or 12 percent agreed, while 118 or 88 percent disagreed with the statement.

Thus, most of the respondents agreed that CME programs can and must be evaluated in terms of their effect in improving patient care rather than by pre- and postprogram paper and pencil exams.

Three further statements on this opinion survey seem worthy of mention. Statement No. 20 -- "Much valuable physician time is wasted in attending current standard CME programs." One-hundred twenty-five CME program producers responded to this statement -- 56 or 45 percent agreeing, and 69 or 55 percent disagreeing. Such a response of CME program producers implies a necessity for change, although I recognize the word "standard" may be open to various interpretations.

Statement No. 13 -- "Behavioral scientists have little to offer physician educators in the design of CME programs." Of 143 responses only 15 or 10 percent were in agreement; 128 or 90 percent disagreed. Increased use of behavioral scientists in design of continuing medical education programs seems indicated by this response.

Finally, Statement No. 15 -- "Incentives like the 'Physician's Recognition Award' of the AMA are likely to have a strong motivating influence in securing physicians' CME program attendance." of 122 responses only 14 or 11 percent were in agreement, while 108 or 89 percent disagreed.

What do these responses mean? In my estimation, they represent a necessary but insufficient condition for progress. As opinions, they represent intellectual agreement. But we are looking for evidence of evaluation activity on the part of CME program producers. Favorable attitudes should be reflected in educator behavior. Behavioral evidence of educator attitudes favorable to CME program evaluation, particularly that which measures its effect on patient care, seems strikingly absent. What is the problem? Unfortunately, I don't have data on educator attitudes toward evaluation except as can be implied or assumed from the almost complete lack of evaluation of current CME programs, particularly as we move from simple head counts as a method of evaluation, through paper and pencil exams, to the effect of CME programs on patient care process, to the final criterion of CME success which is improved patient care and results. For the purposes of this conference may we accept the seemingly obvious fact that there are overwhelmingly negative attitudes toward evaluation of continuing medical education programs by their producers, particularly as we move toward more meaningful types of evaluation. And let us accept as the challenge of this conference the improvement of educator attitudes toward evaluation of continuing medical education.

II. How to Improve Educator Attitudes Toward Evaluation of Continuing Medical Education

To improve educator attitudes toward evaluation of continuing medical education programs or, more precisely, to increase the number of educators who are evaluating the effect of their CME programs in improving patient care, we need to do three things: a) We must improve the conditions surrounding evaluation, b) we must improve the consequences following evaluation, and c) we must provide models; that is, some highly respected producers of continuing medical education programs who are evaluating the effect of their programs in terms of improvement of patient care.[1]

Positive Conditions Surrounding Evaluation

Here is a list of conditions that might exist during the evaluation process that would tend to <u>increase</u> evaluating activity (action, behavior) on the part of CME program producers.

1. The CME producer can demonstrate to himself and to the learner that the learner is progressing (learning).
2. Learner recognizes producer's interest in him in providing evidence of his success.
3. Learner recognizes that evaluation of education in terms of its effect on his practice keeps education relevant to his needs, his patient care problems, his practice.
4. Feedback to learner from evaluative process enhances learning where learner behavior is appropriate (the principle of programed learning).
5. Inappropriate learner behavior is not supported, making <u>un</u>learning (a very difficult process) unnecessary.
6. Saves producer and learner time -- don't teach what learner knows.
7. Makes more individualized teaching possible.
8. Game atmosphere that can exist during the evaluative process can be enjoyable.
9. Increased grant support usually obtained for programs that are well evaluated.

Aversive Conditions Surrounding Evaluation

Here is a list of conditions that might exist during the evaluation process that would tend to decrease evaluating activity (action, behavior) on the part of CME program producers.

1. Demonstration of lack of program success.
2. Demonstration of inappropriateness of learning experiences to achieve agreed upon educational objectives.
3. Producer frustration in achieving skills necessary to evaluate.
4. Evaluation can be expensive.
5. "Loss" of producer and learner time by participating in the evaluative process.
6. Learner fear of producing evidence of lack of achievement, whether evidence is from exams, observation of practice skills, chart review, or any measures that show lack of improvement of patient care process or end results.
7. Learner fear of evaluative process as a carryover from medical school where certifying function of exams may have been stressed rather than their effect of enhancing learning through feedback.

Positive Consequences Following Evaluation

Here is a list of consequences that may follow evaluation activity that would tend to increase furtner evaluating activity (action, behavior) on the part of CME program producers.

1. Educator has solid evidence that the education program was a success -- patient care improved.
2. If the desired behavioral results were achieved it may indicate that the educational techniques, learning experiences, were appropriate to the objective.
3. Evidence of success may improve subsequent physician attendance and participation.
4. Improved patient care directly related to education program activity provides the more powerfully motivating force of an intrinsic incentive.
5. Learning time may be decreased with feedback from the evaluative process.
6. Evidence of success may bring additional financial support.
7. Success in terms of improved patient care may justify the CME producers' existence -- in fact, what else should?

Aversive Consequences Following Evaluation

Here is a list of consequences that may follow evaluation activity that would tend to decrease further evaluating activity (action, behavior) on the part of CME program producers.

1. Educator may secure evidence that no behavioral change occurred in the physician; patient care was not improved; the program was a failure.
2. Failure of the program as documented by evaluation activity may indicate the learning experiences, the education techniques, were inappropriate to the educational objectives.

3. Lack of demonstrable success may decrease attendance.

4. Lack of learner achievement may discourage the learner.

5. Lack of relation of education program activity to patient care improvement may be demonstrated to the learner.

6. Failure may require that the educator learn something about education.

7. "Loss" of learner and educator time in evaluative activity.

8. Evidence of failure may reduce financial support.

Modeling

Finally, to increase the incidence of evaluative behavior on the part of CME program producers, we must have acceptable models of program evaluation exhibited by educators who command respect. Herein lies the second challenge to those present. If most of you begin evaluating your CME, you could provide the models and modeling necessary. Very possibly you have this responsibility to those in the field of CME who accept you as their leaders. We feel that we have an acceptable model at Chestnut Hill Hospital, at least for community hospitals, and through workshops I will be conducting at the Universities of Illinois, Colorado, Washington, and in California, will be attempting a replication of our model. We need acceptable models at other levels; i.e., medical school CME programs and RMP's.

Acceptable models of CME program evaluation will be those that exhibit most of the positive conditions and consequences surrounding and following evaluative activity and few of the aversive conditions and consequences.

Reference

1. Mager, Robert F.: Developing Attitudes Toward Learning. Palo Alto: Fearson Publ., Inc.

Select those statements with which you are in greatest agreement and place the number that preceeds them in these boxes.

( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

1. CME directors usually know the course content needed for physician participants.
2. Success of CME programs must be measured in terms of improved patient care.
3. The central problem in CME program development is motivating the practicing physician to attend.
4. In the design of CME programs, the patient's needs should be met first by defining areas of poor patient care.
5. A CME program is best held in a location where the practicing physician can get away from his practice and its problems.
6. If the care actually provided to a patient population can be compared to an ideal or criterion physician performance, the physician behavior needed to transform actual care to ideal care can be translated directly into education objectives.
7. Most practicing physicians have little desire to keep up with the advances in medical practice.
8. A system of priorities seems essential in determining which patient care needs shall be met first.
9. It is a waste of time to ask physicians what they want before designing CME programs.
10. In deciding where to start a CME program, one should determine the greatest causes of preventable disability and select those where physicians are doing little to prevent disability.
11. Behavioral scientists have little to offer physician educators in the designs of CME programs.
12. It will be absolutely essential to educate substantial numbers of allied health personnel to do many of the things physicians are now doing if we are to meet the health care needs of the country.
13. Incentives like the "Physician's Recognition Award" of the AMA are likely to have a strong motivating influence in securing physicians' CME program attendance.
14. It is possible to measure the effect of CME programs on patient care, and also to determine the need for CME programs by measuring patient care.
15. The true success of a CME program can best be judged by a fore and post-examination (knowledge test) of the participants.
16. Much valuable physician time is wasted in attending current standard CME programs.

Now select those statements above with which you are in greatest disagreement and place the number that precedes them in these boxes.

( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

Now please respond to these statements.

17. I have evaluated at least one of the CME programs I have conducted by measuring its effect on patient care. Yes____ No____

18. I have developed the curriculum for at least one of the CME programs I have conducted by first gathering data to substantiate patient care deficit (s). Yes____ No____

19. I have gathered data representing patient care deficits before and after at least one of the CME programs I have conducted and found significant improvement in patient care after my CME program. Yes____ No____

20. I have been unable to overcome the obstacles present at my hospital to center my CME programs on documented patient care deficits and to measure effects of my program on these deficits. Yes____ No____

21. I feel that my medical staff and/or I do not have the skills necessary to center our CME program on patient care needs (deficits). Yes____ No____

22. My medical staff and I would like to achieve the skills necessary to center our CME programs on patient care needs (deficits). Yes____ No____

PHYSICIAN ATTITUDES TO EVALUATION*

Neal A. Vanselow, M.D.

In spite of recent emphasis placed upon evaluating the effectiveness of continuing medical education programs, it is surprising that a search of the literature reveals little objective data on the attitude of the practicing physician toward such procedures. In preparing this paper, a concerted effort, including a MEDLARS search, failed to reveal much more than a few casual references to this topic. For this reason, the conclusions arrived at here are based as much upon opinion as they are on fact. A prospective study aimed at determining physician attitudes toward the various types of evaluation schemes now in use or proposed for the future seems highly desirable before large-scale evaluation efforts are undertaken.

Three basic assumptions are made in discussing this topic. The first of these is that doctors in practice have had little training in educational theory and educational techniques and are therefore not inherently attuned to the crucial importance of evaluating the effectiveness of an educational undertaking. Unfortunately, the only evaluation most physicians have been exposed to in their undergraduate and graduate training programs is evaluation aimed at testing the knowledge and skills of the student himself. In this atmosphere it was usually assumed that any failure was due to a defect in the student, and never to a defect in the educational experience. Because of this it appears obvious that those who are interested in evaluating continuing medical education programs have a selling job to do. The practicing physician must be convinced that the time and effort he contributes to evaluation will ultimately be of benefit to him.

A second assumption proposes that selling evaluation will not be as difficult as some would have us believe, and for a very important reason -- the practicing physician is genuinely interested in the quality of medical care he delivers. If he can be convinced that evaluation will ultimately result in improved care for his patients, he will be willing to take the time and effort to cooperate. It hardly seems necessary to defend this assumption, but for those who wish to challenge it, the following data are offered. In 1969, the Center for Research on Utilization of Scientific Knowledge at the University of Michigan, in cooperation with the Department of Postgraduate Medicine of the University of Michigan Medical School, conducted a survey of a stratified random sample of physicians practicing in Michigan. This Michigan Physicians' Survey[1] included 1,600 of the 8,000 Doctors of Medicine and Doctors of Osteopathy in the state, and was designed to learn more about their attitudes toward continuing medical education. Usable data was collected from 1,100 physicians (70% of those surveyed). As part of the survey, each participant was asked to indicate the degree to which he was bothered or worried by a number of factors. Table 1 illustrates that, of the "top ten wories" of physicians in Michigan, the first is concern for the quality of work they are doing. In addition, two of the top three factors which bother physicians in practice relate directly to the matter of quality medical care.

The final assumption, if valid, may exert an adverse effect on physician acceptance of evaluation. In recent years there has been increasing pressure from a variety of sources for mandatory re-certification or re-licensure of physicians.[2,3] Data from the Michigan Physicians' Survey (Table 2) illustrates that a majority of doctors in practice are concerned about this trend. Approximately half of the participants responded in the two most negative categories when asked to indicate the extent to which they favored periodic re-examination as a requirement for liscensure. Only 15 percent of the physicians surveyed responded in the two most positive categories. This data is of importance to the medical educator since many of the techniques now used to evaluate the effectiveness of a postgraduate course can also be used for policing or punitive purposes. For example, direct testing of a physician's knowledge or a medical audit, designed to measure the level of patient care he delivers, could also provide data for re-licensure or re-certification. It appears valid to assume that recent pressure for mandatory re-evaluation of a physician's competence will exert a negative influence on his receptivity toward meaningful evaluation of continuing medical education. The

* Prepared by Neal A. Vanselow, M.D. and William L. Schmalgemeier, M.A.

burden will be on the educator to convince his physician-student that the data he collects will be used only for evaluation purposes and will not be released to agencies concerned with policing the profession.

Present Evidence on Physician Attitudes -- Fact and Opinion

Four major evaluation techniques have been used in recent years. They are: the "head count"; the "popularity contest"; measurement of factual knowledge by pretesting and posttesting; and measurement of changes in physician behavior as related to patient care. In the order listed, each technique is generally conceded to be more satisfactory and sophisticated than the preceding one.

The "head count" technique has been used for years on the rather naive assumption that the effectiveness of a postgraduate program can be measured adequately by counting the number of physicians who attend it. Because of the obvious fallacies in this line of reasoning, as well as the fact that it involves no effort on the part of the attendee, no comment need be made regarding its acceptability to the practicing physician.

The "Popularity contest" is the time-honored technique of asking each enrollee for his subjective impressions of each presentation and the educational program as a whole. This technique has been used in our courses at the University of Michigan for years. The evaluation form we use, shown in Figure 1, is probably no better and no worse than forms used elsewhere, and is presented merely as an example of this type of evaluation. In our experience, such evaluation efforts have been well accepted by the physician in practice. The course director must continually remind the course enrollees to fill out the forms but, if such reminders are given, a high yield of completed evaluations is obtained. The resultant comments are usually candid and helpful in planning future activities. We do not require the physician to put his name on the evaluation form. The resultant anonymity has probably been helpful, since the evaluation form is in no way threatening to the individual completing it.

An interesting variation of the "popularity contest" approach has been used by Dr. Thomas C. Meyer[4] at the University of Wisconsin Medical School. Prior to a course, a small number of enrollees is selected randomly and asked to meet with the course director. They are queried regarding their expectations, and are also requested to meet with the course director at the end of each session to critique the presentations given that day. Dr. Meyer has found this technique exceedingly useful. He has observed that the enrollees will give pertinent and candid comments, even when their critiques are taperecorded. In his experience, this type of evaluation technique has been eagerly and enthusiastically accepted by the practicing physician.

Although the evidence suggests that the "popularity contest" approach to evaluation is well accepted by physicians in practice, the evidence for acceptability is less clear for techniques which include testing of factual knowledge. In the latter, each physician-student is tested on course content before and at varying periods after the course. Short objective tests of the multiple choice variety are commonly used. This method of evaluation is potentially more threatening to the practicing physician, since it permits a third party to assess directly what he knows and what he does not know. In spite of this, Dr. Stephen Abrahamson has stated that, in his experience, short objective tests administered at the beginning of a program and at its completion seem to be routinely accepted by the participants.[5]

Two other bits of evidence would seem to confirm this view. The first of these is experience with the Self-Assessment Test of the American College of Physicians.[6] Developed by the College's Committee on Educational Activities, this test permits practicing internists to test their knowledge in nine areas of internal medicine by voluntarily agreeing to take a 700-question multiple-choice examination at home. Initially, 5,000 of the 15,000 ACP members took the test. Since the initial offering, an additional 5,600 members and nonmembers have participated. From all indications,[6,7,8] the program has been enthusiastically received, and the American Psychiatric Association and American College of Pediatrics are planning similar undertakings. Caution must be used however

when attempting to generalize from this experience. The Self-Assessment Test was accepted by practicing internists, but great care was taken to minimize its potentially threatening aspects. It was purely voluntary. A bonded firm (separate from the College) handled the applications, mailed the test papers, and scored the test results. Participating physicians could be identified only by coded numbers. Only the participating physician had access to his performance. Because of the fear that the data might be improperly used, no group scores were compiled in spite of the fact that such data might have been valuable in directing future continuing education efforts of the College. Under these circumstances, internists cooperated; what might have happened if confidentiality had not been guarded so closely is a matter for speculation.

The response to the Individual Physician Profile project at the University of Wisconsin also suggests that physicians will permit testing of their factual knowledge when done in a discrete and nonthreatening way.[4] Developed by Dr. Thomas C. Meyer and associates, this program was initiated in 1968.[4] Medical secretaries visited the office of each participating physician and recorded data on the type of patient seen in the office, at home, and in the hospital. Based upon his practice profile, a multiple - choice examination was constructed for each physician, testing his knowledge in those areas of medicine which were commonly encountered in his practice. From his practice profile and performance on the examination, a program of continuing medical education was then individually tailored for each participant. Thirty-seven physicians participated in the first year of the study and responded enthusiastically. Only one of the original 37 declined to participate a second year. The original experience may be misleading however, since most of those participating were friends or personal acquaintances of Dr. Meyer. In addition the original group was a highly motivated one, already participating in continuing medical education and considered to be secure in the quality of medical care they were providing. Of more significance is the fact that 60 physicians are participating in the second year of the project and that the number of volunteers has exceeded the number of places available. While this project is an experimental one and is not strictly analogous to the situation of testing knowledge for evaluation purposes, it does suggest that physicians are willing to submit to tests of factual knowledge when used for what they perceive to be constructive purposes.

The most meaningful type of evaluation is felt to be measurement of changes in physician behavior which occur following an educational program. Such measurements can be made most easily in a hospital setting and most commonly utilize the medical audit as a measurement device. The classical studies of Williamson et al.[9] at Rockford Hospital and the well-known program of Brown and associates[10] at Chestnut Hill Hospital in Philadelphia are excellent examples of what can be accomplished when the medical audit is used to design and evaluate programs of continuing medical education. While the medical audit was accepted by the staffs of both hospitals these two studies are somewhat unique, since each was initiated by an outstanding and persuasive individual. One could argue that each succeeded largely because of the personal charisma of the individual in charge, and that it would be dangerous to assume equal success could be anticipated in other settings.

The Medical Audit Program (MAP) of the Commission on Professional and Hospital Activities (CPHA) is used by over 1,300 community hospitals in the United States and several foreign countries. The director of CPHA, Dr. Vergil Slee[11] has found that acceptance of the medical audit is dependent in large part upon the individual who directs it in the hospital involved. In his experience, some physicians have committed "professional suicide" because they failed to take into consideration the worries of the medical staff regarding the potential consequences of a program of this type. Dr. Slee feels the fear of personal embarrassment is the biggest single factor contributing to physician reluctance to cooperate. The fear of third-party sanctions and the feeling that coercion is involved are also given as reasons for opposing the medical audit. Finally, some physicians object to the technique itself. "All you are measuring is medical work," is another. However, CPHA has also found that most medical staffs perceive their original fears to be unfounded once MAP is in operation.

Under a Regional Medical Program grant, Dr. Joseph Hess at Wayne State University Medical School has recently instituted a program of continuing medical education in three small inner-city Detroit hospitals. A medical audit is used to determine educational needs and evaluate the effectiveness of the programs instituted to meet them. Dr. Hess[12] has found each hospital staff willing to accept a medical audit, but emphasizes that two conditions must be met. First of all, the audit must be conducted within the existing staff structure of the hospital; in his experience there has been great reluctance to permit an "outside agency" to come in to look at medical records. Secondly, the audit must be performed and used in a nonthreatening, nonpunitive way. The latter theme constantly recurs in discussing this level of evaluation with knowledgeable individuals. Its importance was perhaps best summarized by Bernard Dryer when he wrote: "What is required then is a clear-cut and widely understood distinction between education and regulation; so that the energy available for learning can be efficiently utilized, with each individual physician aware that only he will decide when and where he will approach a regulatory body for the more formal awards they offer."[13]

Additional Data on Physician Attitudes

In an attempt to obtain some prospective data on physician attitudes toward evaluation, we recently surveyed the opinions of the enrollees in our annual Northern Michigan Summer Course. This course is presented in a resort area each year and is aimed at the family practitioner. Five half-day sessions are devoted to a discussion of a broad variety of clinical topics by carefully selected University of Michigan Medical School faculty members. In order to give the participants some experience with at least one method of evaluation, a short multiple-choice pre-test on the material to be covered was administered on the first morning of the course. A posttest of similar format was given on the final day.

Each participant was then asked to fill out a questionnaire designed to determine his attitudes toward a variety of evaluation technics. Forty of the 52 full-time registrants in the course (77 percent) completed the questionnaire. The results of this study are summarized in Figures 2-5.

Figure 2 illustrates that the great majority of physicians rejected the idea that it was either unnecessary or impossible to evaluate the effectiveness of a postgraduate program. Few felt that counting attendance was an adequate method of evaluation. Measurement of subjective reactions to the program and testing of factual knowledge by pre- and post-testing were rated by most physicians as appropriate types of evaluation. Less than half of those responding felt that evaluation of physician behavior via a medical audit was appropriate, which indicates that attempts to utilize this technique might be met with considerable physician resistance.

The data in Figure 3 confirms the postulate that practicing physicians are more receptive to tests of factual knowledge than they are to evaluation techniques employing the medical audit. The negative response to identification of participating physicians by name suggests that the extent to which anonymity can be preserved is an important factor in determining physician acceptance of evaluation.

Figure 4 demonstrates that those responding recognized that evaluation of continuing medical education might ultimately result in improved patient care. The suggestion that testing a physician's knowledge was an irrelevant nuisance was rejected. The importance of providing each physician with the results of his performance was strongly emphasized. A mixed response was obtained to items b, d and f, each of which probed for possible concern over perceived "misuse" of evaluation data.

The question illustrated in Figure 5 was designed to determine how much physician anxiety might be generated by evaluation techniques designed to measure factual knowledge or physician behavior in delivering patient care. Surprisingly, the response to this item indicated less anxiety than might have been predicted, and evaluation methods which measured behavioral change did not appear to differ significantly from those which measured factual knowledge.

Each item in the questionnaire was analyzed to determine if age or performance on the pre-test and posttest might influence physician response. The responses of physicians over 45 were no different than those 45 and under; however, as might be expected, physicians showing low gain in knowledge from the course tended to be less receptive to knowledge testing as a technique of evaluation.

Summary

1. There is little published data on practicing physicians' attitudes toward evaluation of continuing medical education programs. Prospective studies are needed to provide the medical educator with the information he requires to design evaluation techniques which will be acceptable to doctors in practice.

2. Present evidence suggests that practicing physicians consider the "head count" to be a superficial and relatively meaningless evaluation technique.

3. Surveys of subjective participant reaction to an educational program are regarded favorably.

4. Techniques designed to measure change in factual knowledge are more acceptable than techniques such as the medical audit which are designed to measure alterations in physician behavior related to patient care. Both are perceived as posing a potential threat to the physician in practice, but there is some evidence to indicate that they can be employed successfully, provided:

   a. An effort is made to explain what is being done and the importance of doing it.
   b. The physician is provided with feedback on his performance in a constructive and nonthreatening manner.
   c. Knowledge or performance testing for purposes of evaluating a continuing medical education program is clearly separated from similar testing for punitive or regulatory purposes.

Table 1

Michigan Physicians' Survey

"Top Ten Worries" of Practicing Physicians

"To what extent are you bothered or worried by any of the following things?"

Item	% Responding to Same, a Great, or a Very Great Extent	
	M.D.	D.O.
How good a job I am doing	62	75
The effective use of my time	46	53
Feeling that I am not keeping up to date	43	56
Having to do things I don't want to do	42	41
Having to put up with and tolerate incompetence	41	50
Being torn by conflicting demands	33	31
Failure to advance professionally	32	40
Feeling I am "in a rut" professionally	30	41
Failure to get ahead financially	28	35
Can't do what is expected of me	28	38
	N=919	N=183

Table 2

Michigan Physicians' Survey

"To what extent do you favor periodic re-examination which would be required to maintain your license to practice medicine?"

Percent Responding	M.D.	D.O.
To a very little extent	33	35
To a little extent	16 (49)	13 (48)
To some extent	26	28
To a great extent	11	10
To a very great extent	4 (15)	6 (16)
No response	10	8
	N=919	N=183

Figure 1

E V A L U A T I O N

Department of Postgraduate Medicine

Physiology for the Clinician:
A Practical Review of Selected Topics
February 16-18, 1970

These evaluations are of great help to our course chairman and faculty in planning future continuing education programs. Please be frank.

I. Area of Practice (Generalist, Internist, other - please specify):

II. Graduate of_____ Year_____

III. Please evaluate each session of the course:

	General Evaluation Excellent	Good	Poor	Was it of Practical Value to You?	Comments
Monday, February 16 Review of Basic Mechanisms of Gastrointestinal Control					
Current Methods of Evaluating Liver Function					
Autoimmune Mechanisms in Intestinal Disorders					
Pathogenesis of Acute and Chronic Pancreatitis					
Respiratory Mechanics					
Pulmonary Acid-Base Problems					
Clinical Patterns of Disordered Pulmonary Function					
Use of Respirators in Clinical Practice					
Tuesday, February 17 Basic Mechanisms of Renal Transport					
Clinical Evaluation of Renal Function					
Mechanisms of Uremic Manifestations					
The Nature of Renal Tubular Acidosis					
Methods and Measurements of Dialysis					
Normal Physiology of the Renin-Angiotensin System					
Control of Aldosterone and Sodium Metabolism					
Secondary Aldosteronism Without Hypertension					

Figure 1 (continued)

	General Evaluation Excellent	Good	Poor	Was it of Practical Value to You?	Comments
Film: Aldosterone					
Basic Mechanisms of Myocardial Control					
The Pathophysiology of Myocardial Failure 1. In Ischemic Heart Disease					
2. In Valvular Heart Disease					
3. In Congenital Heart Disease					
Basic Mechanisms of Peripheral Vascular Control					
The Pathophysiology of Occlusive Vascular Disease					
Televised Demonstration: Methods of Graded Exercise Testing					
Hemodynamic and Electrocardiographic Responses to Exercise					

IV. The outstanding feature of this course was:

V. Suggested improvements before this course is offered again:

VI. Other comments (use back of sheet if necessary):

Figure 3

Practicing Physicians' Reactions to Evaluation Techniques

Northern Michigan Summer Course - June, 1970

a.	Testing factual knowledge at beginning and end of postgraduate course.	Strongly Agree	4	3 V	2	1 Strongly Disagree
b.	Testing factual knowledge three or six months after attending course.	Strongly Agree	4	3	V 2	1 Strongly Disagree
c.	Responding to test with name for identification.	Strongly Agree	4	3	V 2	1 Strongly Disagree
d.	Medical audit with intent of improving practice.	Strongly Agree	4	3	V 2	1 Strongly Disagree
e.	Test by national specialty or subspecialty organization.	Strongly Agree	4	3	V 2	1 Strongly Disagree

N=40

Figure 4

Practicing Physicians' Reactions to Evaluation Concepts

Northern Michigan Summer Course - June, 1970

a.	Testing a physician's knowledge with feedback of performance is worthwhile.	Strongly Agree	4	3 V	2	1 Strongly Disagree
b.	Testing a physician's knowledge is worthwhile to others.	Strongly Agree	4	3	V 2	1 Strongly Disagree
c.	Direct testing of a physician's knowledge is a nuisance.	Strongly Agree	4	3	V 2	1 Strongly Disagree
d.	Testing of a physician's knowledge is a first step toward recertification.	Strongly Agree	4	3	V 2	1 Strongly Disagree
e.	Testing a physician's knowledge is good for educational purposes but doesn't help his patients.	Strongly Agree	4	3	V 2	1 Strongly Disagree
f.	Maintenance of a licensed physician's competence is his personal concern.	Strongly Agree	4	3	V 2	1 Strongly Disagree

N=40

FIGURE 2

NORTHERN MICHIGAN SUMMER COURSE-JUNE 1970

Physicians' Selections of Appropriate Postgraduate Course Evaluation Techniques

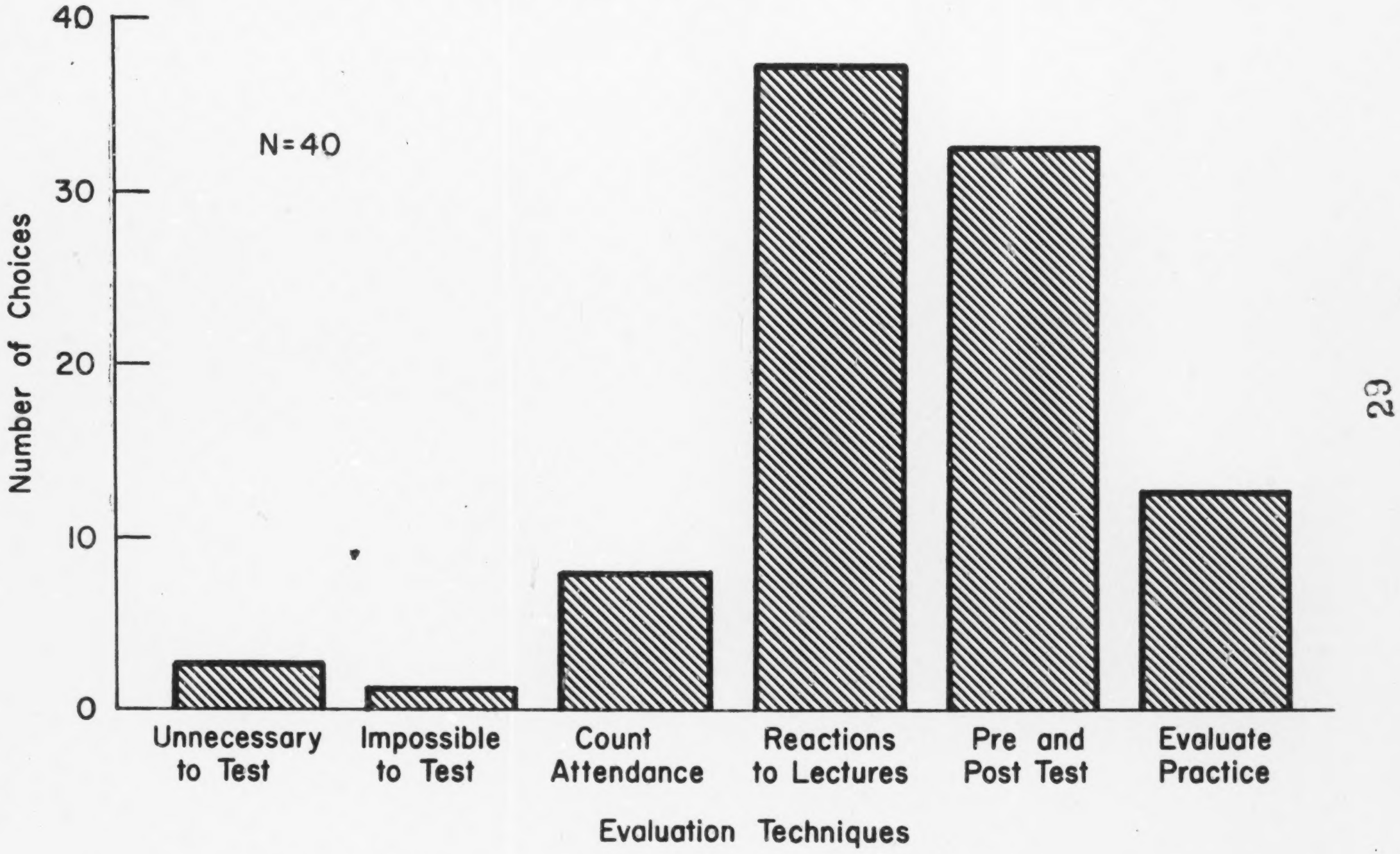
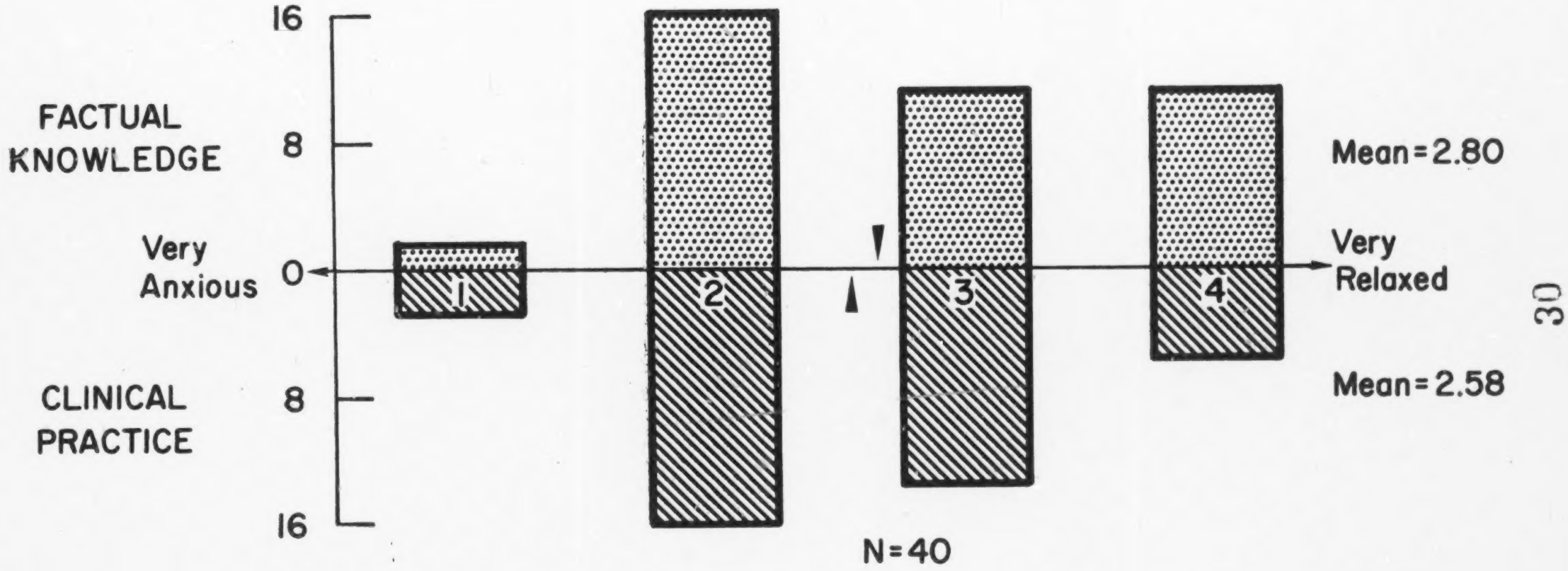


FIGURE 5

NORTHERN MICHIGAN SUMMER COURSE – JUNE 1970

Physicians' Attitudes to Being Tested



References

1. Mann, F. C.; Reilly, A. J.; and Vanselow, N.A.: The Michigan Physicians' Survey (in preparation).

2. Millis, J. S.: The Impact of Re-Licensure Upon Postgraduate Medical Education. Mod. Med., June 29, 1970.

3. Trimpi, H. D.; Martin, P. V.; and Khubchandani, I. T.: Compulsory Continuing Physician Education in the United States. Lex et Scientia, July-September, 1969.

4. Meyer, T. C.: Personal communication.

5. Abrahamson, S.: Evaluation in Continuing Medical Education. J.A.M.A. 206:625, 1968.

6. Rosenow, E. C., Jr.: The Medical Knowledge Self-Assessment Program. J. Med. Educ. 44:706, 1969.

7. The Self-Assessment Test (Editorial). J.A.M.A. 209:764, 1969.

8. Self Assessment of Medical Knowledge (Editorial). New England J. Med. 278:964, 1968.

9. Williamson, J. E.; Alexander, M.; and Miller, G. E.: Priorities in Patient-Care Research and Continuing Medical Education. J.A.M.A. 204:303, 1968.

10. Brown, C. R.: Assessment of the Quality of Patient Care (presented at 66th Annual A.M.A. Congress on Medical Education, February 9, 1970).

11. Slee, V. N.: Personal communication.

12. Hess, J.: Personal communication.

13. Dryer, B. V.: Lifetime Learning for Physicians -- Principles, Practices, Proposals. (Report from the Joint Study Committee on Continuing Medical Education.) J. Med. Educ. 37:81 (Part II), 1962.

INAPPROPRIATE OR UNREALISTIC OBJECTIVES

Stephen Abrahamson, Ph.D.

The classic paradigm in education (and educational evaluation) has been preached (and practiced) for some years. Attention paid to this paradigm has sharply increased over the last several years as "educationists" have become more deeply involved in medical education and in continuing medical education. Almost a litany has emerged -- a powerful hymn, the opening line of which (while it may have variations) basically goes like this: "First, you must define your objectives."

The remainder of the paradigm, of course, suggests that only then can one plan the learning experiences, the use of audiovisual aids, the necessary measurement procedures, and the like. For those of you still unfamiliar with this theme and its variations, there are numerous sources to inform and infect you: Miller's Teaching and Learning in Medical Schools[1]; Abrahamson's article in JAMA, "Evaluation in Continuing Medical Education"[2]; Mager's Preparing Instructional Objectives.[3] The term "infection," is used because there are interesting signs and symptoms which appear in one who has first been "exposed" to this bit of learning (whether he learns it or not).

The most dangerous reaction to this apparently harmless paradigm is internal and may manifest itself in a variety of ways. The person exposed seems to assume that following the steps in educational planning and/or educational evaluation is as simple as the statement of those steps. Perhaps this is a tribute to those who have "preached the gospel" and won new followers, for we "preachers" have deliberately set out to make the processes seem logical, necessary, and worth pursuing. What we have neglected, for the most part, is to warn that this road to "educational Heaven" is a difficult and deceptive one. It was not for some time, for instance, that "preachers" began attempting to practice what they preached -- at least in the setting of medical education and continuing medical education. Rosinski's early experience at the Medical College of Virginia in 1959 may be the first documented and significant involvement of an educationist in the formal statement of real educational objectives in medical education.[4] And the biggest assistance in this area came from Mager in 1962, when he wrote a little programmed text (then called "scrambled text") on how to write objectives to guide the preparation of programmed instruction.[5]

Mager's most significant contribution has been said before by many -- but never so well, apparently, at least to judge by the impact. Indeed, Wrinkle said it in 1947 in a list of "Criteria for Educational Objectives," in which he suggested that educational objectives must be stated "in terms of the learner's behavior."[6] By this he meant what Mager has since promulgated in exquisite detail. Educational objectives are desired outcomes -- changes in what the learner does (or demonstrates he can do) as a result of the learning experience. In fact, "as a result of" may be only my inadvertent addition (a Freudian slip, perhaps?) and should read only "after" the learning experience. It may be only the educational scientist who has the deep interest in such things as cause-effect relationships.

All of this discussion should point out the most glaring inappropriateness in educational objectives -- one which hardly needs mentioning, let alone discussion. It is now generally accepted that educational objectives should describe what the learners are doing (or can demonstrate that they can do) when the instruction is completed. It is inappropriate to include the "reasons" for holding a program; e.g., to "cover" the new biochemistry; to "reach" the men "in the field" or "in the front lines"; to "demonstrate" the new techniques; to "report" the latest "word"; to hold our annual symposium. (In all too many instances, that last example is the only statement of objectives to be found -- and to think: it is now declared "inappropriate.")

Of course, it is possible to define these desired outcomes at different "levels." Jack Thomson has done this in his distinction between "goals" and "objectives."[7] Miller and his associates did this in making a distinction among "ultimate," "intermediate," and "immediate" objectives.[8] Those planning programs of continuing medical education may

want to refer to what the physician-participants will learn in the program; they may want 21
to refer to how the physician-participants will change after the program; they may want to refer to how health-care patterns or health status itself may be affected. Thus, objectives may be defined as desired outcomes at several levels.

1. The physician-participant will acquire knowledge, gain skills, and/or develop attitudes to certain designated or defined standards.

2. The physician-participant's performance in certain areas will be "improved" in certain designated ways.

3. The physician-participant's professional practice will be affected in designated directions.

4. Health-care practices in institutions and communities of physician-participants will be modified in designated ways.

5. The health of communities of physician-participants will be significantly better in designated areas.

"Realistic" objectives, then, can be defined as those which meet this single criterion: there should be reasonable expectation of achievement of the objectives as a result of the program. A four-day symposium on early detection of respiratory disease through the use of spirometry, diagnostic radiology, and other screening procedures -- a symposium which includes lab sessions for physician-participants to use these diagnostic procedures under supervision and which offers extensive case reviews -- may be expected to improve physician-participants' skills in this area. The symposium might even be expected (reasonably) to bring about significant changes in professional performance and practice of physician-participants. The key phrase "reasonable expectation of achievement" will be illustrated further later.

The definition of "appropriate" objectives requires an object to the phrase: "appropriate to..." In continuing medical education, objectives may be said to be "appropriate" when they are appropriate to the health needs of the physician-participants' respective communities and/or institutions, to the practice needs of the physician-participants, and/or to the educational needs of the profession. In a hospital where medical audit has revealed relatively poor management of patients who have suffered a stroke, it would be appropriate to offer a program designed to help physician-participants "improve" their practice in this area. In a community in which there has been a sharp increase in the incidence of venereal disease, it would be appropriate to offer a program designed to help physician-participants learn about community resources available for assistance to patients in this area. For physicians who "feel" or demonstrate certain lack of information (or skill) it would be appropriate to offer a program designed to help meet these felt needs. Finally, lest the author be labeled anti-science or anti-intellectual, in instances it might be appropriate to offer a program designed to help physician-participants acquire some new basic (or basic science) information now thought to be relevant and/or important to the practice of medicine. (Note the qualifiers in that sentence; they will need clarification and defense later.)

Perhaps the best way to further distinguish between "realistic" and "unrealistic," and between "appropriate" and "inappropriate" is to present some examples. And, as in the case of defining "normal" by illustrations of the "abnormal," it might be fruitful to describe what this author believes to be inappropriate and/or unrealistic, especially since this is a "position" paper.

An exaggerated illustration, for instance, of inappropriate objectives is drawn from an emerging nation in Africa. Designing a continuing medical education program with the (correctly, even) stated objective of having physician-participants learn skills of vector cardiography -- when there are many significant and unmet health needs not including heart disease -- seems to this author to be inappropriate....inappropriate to the health needs of the community, inappropriate to an extreme in this regard. Surely, time would be better spent attempting to have physician-participants learn in areas which

might significantly affect (for the better) the health of their respective communities. And our affluence in the United States does not render us immune to the need to apply this kind of thinking in planning continuing medical education. Your own experience will provide many examples of programs with objectives completely unrelated to the health needs of the communities and institutions of participating physicians -- even in the face of serious health problems within those institutions and communities. The work of John Williamson, Bob Evans and Clem Brown surely substantiates this position.

Objectives may be inappropriate to the practice needs of the physician-participants. When this happens, one may find what the American Heart Association reported in a now familiar, classic, and (unfortunately) all-too-often disregarded study.[9] Using a pre-test, post-test and delayed post-test design, they discovered a significant improvement in physician-participants' ability to recognize arrhythmias as a result of an intensive training program, only to see those "gains" disappear after a period of time -- during which those physicians undoubtedly had no occasion to apply those new-found skills. Thus, one can say that the objectives of the original program were inappropriate....inappropriate to the practice needs of those physicians.

Another example of this same kind of inappropriateness is taken from a study reported by Judilynn Foster and Sandra Lass.[10] While their purpose in conducting that study was not what is mentioned here, they serendipitously discovered something germane to this paper, and very interesting. They were attempting to evaluate the effects of a videotape on the practices of physicians in a hospital.....in this instance, ordering serum-sodium and serum-potassium determinations and using potassium supplement where indicated with patients who were on diuretics. Two "matched" hospitals were employed in the study: in one of them, records were reviewed prior to and after the presentation of the videotape; in the other, similar record review took place at identical times but without the "educational treatment" of the videotape. Interestingly enough, Foster and Lass discovered that physicians in the "experimental" hospital (i.e., the one in which the videotape was shown) were already "performing" at a high quality level of practice (with regard to the care needed with patients on diuretics) prior to the introduction of the videotape. Thus, one can say that the objectives for using that program were inappropriate....inappropriate certainly to the practice needs of those physicians.

The "saving grace" for most continuing medical education programs -- at least with regard to appropriateness of objectives -- comes with the inclusion of the modifying phrase "to the educational needs of the profession." Any profession has responsibility for constantly monitoring the quality of service performed and for keeping practitioners performing at desired quality levels. Thus, the medical profession, through its agencies of continuing medical education, has the responsibility for helping its practicing members "keep up" by acquiring new knowledge, by gaining new skills, by reinforcing attitudes of self-examination, self-education, and self-improvement. And, it is the professions' responsibility to define what practitioners should know and know how to do. But it is the responsibility of the profession and not that of individual scientists or of individual clinicians to do so. The late Chester Hyman, Professor of Physiology at the University of Southern California, reported on an interesting incident in his professional life. He was asked to participate in a continuing education program and present a series of lectures on the topic, "What's New In Physiology Since You Left Medical School?" He asked the hundred-odd physician-participants to indicate when they had graduated from medical school. The responses ranged from three years before the symposium to fifty-three years before it! The point here is that the profession had abdicated to one physiologist the determination of what was "appropriate to the educational needs of the profession."

It is, perhaps, in this last area of appropriateness that the author has a "position" which may really stir up some controversy. Determination of "educational needs of the profession" should not be left to the discretion of individuals, however expert in their own fields. If possible, these experts should work with those who are planning programs as well as with representatives of those who will be the physician-participants. In that way, their own intensive studies and highly specialized knowledge will be best applied to the community health needs, the institutional health-care needs, and the individual practitioner's educational needs. In cases where it is impossible to

so involve the expert (for the many reasons known to all of us), it would seem important to do similar kinds of planning with a local -- albeit less well qualified -- "expert" and indicate to the "guest expert" the objectives, the limitations of desired "content coverage," and the like....as tactfully, of course, as possible. Through this approach, it might be possible to retain "control" of both the objectives and the "content." 23

So far, we have talked about inappropriateness of objectives -- to the health needs of institution and community, to the practice needs of the physician, to the educational needs of the profession. Let us consider the matter, now, of unrealistic objectives. To do so, one must also consider the limitations of the program: the duration of the teaching-learning experiences, the size of the learning group or groups, the modes of instruction available, the setting for learning, the characteristics of the learners, the teaching skills of the instructors, and the like. With these factors we must also consider some tried-and-true principles of teaching and learning.

For it is an unrealistic objective for a one-hour lecture on vector cardiography presented to one-hundred physicians to have those physician-participants gain appreciable skill in vector cardiography, as evidenced in their ability to interpret a set of "unknowns." It exceeds reasonable expectation of achievement (thus, it is unrealistic) because one does not gain new skills or improve old ones by listening to a lecture (at least, not usually). It exceeds reasonable expectation of achievement because the amount to be learned demands far more than the time allotted.

It is equally unrealistic to expect physician-participants' habits to change as the result of a lecture or two, however brilliantly delivered. Habits take time to form; they take more time to change. Thus, it is unrealistic for a program to list an objective of establishing certain new habits without the time and method planned which might at least hold some theoretic hope of achievement. Realistically, to change habits would demand at least plenty of time, a series of reinforcing experiences in the new behavior area, small-group discussion of insecurities engendered during the course of learning, and follow-up reinforcement for the new behavior patterns.

Another instance of unrealism in objectives involves misjudgment (or lack of consideration) of the motivation of the learner. To present a videotape to a group of casual "drop-in" learners and hypothesize (another way to state an objective) that they will change their professional behavior significantly is unrealistic. Again, to bring about significant change in behavior requires more than a videotape film, or synchronized slide-tape presentation. It would not be unrealistic, if those modes were used in the ways described, to expect change in attitudes or in levels of professional information within that topical area.

One more example of lack of realism is that of stating objectives of significant change in physician-participants despite a full awareness that the program is being held "at home" for the participating physicians, and that past experience has shown that less than half of those who are in attendance at the outset are still in attendance toward the end, and that most of the physician-participants will be interrupted by calls from the office, the hospital, or the exchange. This is not to say that programs cannot be held "at home"; it does suggest, however, that realistic objectives need to be set.

The importance of setting realistic objectives may be considered far greater when one more factor of appropriateness is introduced. Earlier, appropriateness was limited to three major areas: appropriate to health needs of institution and community, to practice needs of the physician-participants, and to educational needs of the profession. Now it is important to add one more factor: cost. One must ask the question, "Is this objective worth achievement -- at the projected costs in money, personnel time, learner commitment, necessary materials and equipment?"

A few years ago J. S. Denson and I directed a project supported by the United States Office of Education. The purpose (another word for objective?) was <u>to test the feasibility of constructing</u> a plastic-skinned, computer-controlled, anthropometric manikin to be used in training anesthesiology residents. The project was successful, as you know from reading about it in <u>Time</u>, <u>Life</u>, and other education journals, in doing

exactly that: demonstrating the feasibility of constructing a patient simulator which might be used in training health care personnel. Critics (some of them within the bureaucracy of the granting agency) expressed disappointment, saying that we had spent more than a quarter of a million dollars to train five anesthesiology residents -- and only in endotracheal intubation. Had our objective been to train health care personnel, the criticism would have been justified. Had our objective been to train five anesthesiology residents in endotracheal intubation, the criticism should have been that the objective was inappropriate. The objective (when properly stated) might then have been appropriate -- 1) to the health needs of the community, 2) to the practice needs of the physicians, and 3) to the educational needs of the profession -- but still inappropriate, in that an inordinate amount of money had to be committed to a relatively small number of learners to have them master a rather circumscribed skill. For those who see an extended period of time committed by each practitioner to a return to the teaching hospital in a refresher-residency as the only means of "real" continuing medical education, this factor of appropriateness looms of critical importance.

Less starkly, it is possible for large commitments of time by busy physician-instructors and/or equally busy physician-students to be devoted to a teaching-learning exercise, the objective of which is not important enough to warrant such commitment and, thus, is inappropriate.

A "position paper" connotes that the author will assume a position with regard to a given subject -- usually for purposes of stimulating thought and discussion....preferably in that order. In the event that the position is not clear, here are some summary statements.

1. Many continuing education programs are based upon objectives (stated or unstated) that are unrealistic, in that there is not a reasonable expectation of their being achieved.

2. The unrealistic quality comes from a lack of consideration of principles of teaching and learning in the planning.

3. Many continuing education programs are based upon objectives (stated or unstated) that are inappropriate to the needs of the community, the physicians' practice or the profession's needs, and/or are inappropriate because of the inordinate costs involved in light of the gains to be anticipated.

It is further my contention, however, that many programs are planned -- still -- without objectives. These "objectives," of course, are the most innappropriate of all. More than that, with the present fiscal climate, it is frankly unrealistic for programs not to have carefully thought-out and equally carefully stated objectives. During the last several years, sponsors of continuing medical education efforts have begun to ask for evaluative data to answer the question: "Did our money buy anything in the way of improvement?" That question is becoming more of a demand now. The caution is to those of us who are planning programs: make sure that objectives are realistic and appropriate. It is a certainty that we will be held to our "promises."

References

1. Miller, et al.: Teaching and Learning in Medical School. Cambridge: Harvard University Press, 1962.
2. Abrahamson, S.: "Evaluation in Continuing Medical Education." JAMA, 206:625-28, 1968.
3. Mager, R. F.: Preparing Objectives for Programmed Instruction. San Francisco: Fearson Publ., 1961.
4. Rosinski, E. F.: "An Approach to Medical Education: The Development of the M.C.V. Curriculum." M.C.V. Quarterly, 1:3-6, 1965.
5. Mager, op. cit.
6. Miller, op. cit.

7. Thomson, J. E.: Evaluating the Effect of Operational Projects in Regional Medical Programs. University of Southern California, Division of Research in Medical Education. June, 1969.
8. Miller, op. cit.
9. McGuire, C., et al., "Auscultatory Skill: Gain and Retention After Intensive Instruction." J. Med. Educ. 39:120-31, 1964.
10. Foster, J.; and Lass, S.: A Feasibility Study Using Patient Records to Evaluate Videotape Postgraduate Education. A paper presented at the eighth annual conference on Research in Medical Education, Association of American Medical Colleges Conference. October, 1969.

COSTS OF EVALUATION

William T. Herzog, M.P.H.

The primary intent of this paper is to raise issues related to the costs of evaluating continuing education efforts. What are the probable costs of various evaluation efforts? What general resource allocation issues are facing the field of continuing medical education and what role does evaluation play in these issues?

The text of the paper offers some general interpretation of the increasing interest in evaluation and the relevance of cost analysis to this trend. The Appendix estimates the cost of different evaluation approaches of a hypothetical short course in order to provide some basis for comparison. The final paragraph of this Appendix perhaps best states the major resource allocation issue in question.

In the final analysis one cannot escape a basic dilemma common to all human service persuasions in our mechanized culture. On one hand, society is asked to finance extensive educational efforts with very little documentation. On the other, if it tries to correct this by research and analysis, it ultimately must face financing extensive evaluative studies with very little documentation as to the value of the information obtained. The obvious recourse in the face of these uncertainties is for society to give greater attention to the tangible and more measurable options. The only practical way to prevent this is to attempt to develop better estimates of program effectiveness, to argue the intrinsic value of the more qualitative and less measurable aspects of human endeavor, and to bring the best minds to bear on questioning the apparent but deceptive value of the more tangible and visible alternatives. If there is any success, high-quality human service might at some time attain the political and social appeal now reserved for our massive highway systems, space programs, defense capabilities, and impressive building programs.

Implications of Program Budgeting

Although the problem of competing with bricks and mortar is not new, it has taken on some new dimensions that are essential to understanding the cost issue. Beginning on August 25, 1965, a relatively novel concept was introduced into the federal budget system.[1] Alternatively known as program budgeting, PPBS, (Planning, Programming, and Budgeting System), and management by objectives, the concept stemmed from management techniques introduced in the Department of Defense during the early 1960's. The historical roots of program budgeting are found in Frederick Taylor's Scientific Management and are closely linked with the operations research approach developed during World War II and subsequently applied in numerous industrial settings. The fundamental proposition of program budgeting is that allocation of funds for a given program should be directly related to the expected accomplishments of that program when viewed in the context of over-all organizational priorities. This approach to budgeting may sound traditional and unobjectionable, but in fact it represents a major change. Even now the single most important determinant of the annual budget allocation is last year's expenditure pattern.[1,2]

The program budgeting system requires that organizational requests for funds be accompanied by the specific objectives to be attained within the budgetary time frame. The stated objectives become the guidelines for measuring program effectiveness. The ratio of cost to effectiveness becomes a method for evaluating organizational efficiency in order to allocate the limited available funds. Budget review officers are increasingly asking organizations to outline their expected accomplishments, describe the alternative approaches available, and specify which approach is most likely to maximize gain while minimizing the cost and risk involved. The ultimate application of this approach can be found in the business setting where the profit margin is an exact statement of the cost effectiveness ratio.

It is not difficult to relate the current pressure for evaluation to this larger trend in business and government management. At a recent conference on medical continuing education, one of the participants indicated the federal trend by stating, "....There are

many hopeful signs for improvement of (continuing education) programming, because I think the mood now is definitely to look for accomplishments and importance, to look at the objectives. What do you really plan to accomplish with X amount of staff and X amount of money and X amount of materials?"[5] One need only to look at grant guidelines, talk to foundation officials, or in North Carolina to review the new budget request procedures to see the application of this approach and estimate its implications to planning and development.

There is much to be optimistic about in this trend toward more rationalized decision-making. For one thing, it will place a needed spotlight on programs which have not had the judgment and introspection to attempt to measure their achievements and identify their shortcomings. On the other hand, there is good reason to be cautious and discriminating in what we buy from the new management science. The greatest danger is in being overwhelmed by the complexity and to escape either by embracing any method with numerical claims or by running at the first hint of scientific decision-making. Powerful analytical tools are being developed which can be valuable to continuing education, but it remains a world in which the saying caveat emptor applies as well as ever. Application of these techniques to education involves serious methodological and conceptual difficulties, "once (as one author put it) you get beyond the level of easy majestic generalities."[3] While taking the side of objective, rational investigation, systems analysis involves judgment in every stage from problem definition through model building to estimation of parameters.[4]

Evaluation Objectives

Investment in evaluation should follow the same rules we should apply to any planning effort. The first question should be: "What will this study accomplish and what difference will it make to decision making?" In my own experience there seems to be four possible objectives in evaluating continuing education programs: 1) to improve course design, 2) to improve teaching methodology, 3) to contribute to knowledge, and 4) to justify the organizational or societal value of the effort. Some writers in the field of evaluation specify that objectives should be attainable and measurable. I would not force evaluators to undergo the same test, although I would urge that the objectives be clearly stated, meaningful, and as attainable and measurable as possible. Insistence on stating the attainable when you really want to go further is to reduce your true aim. Insistence on quantification when you are dealing with an intangible is to settle for what may be an imprecise attribute of the real object in question. Although it is easy to understand why such mandates are included in the evaluation literature, they have two drawbacks: 1) they can tend to mask the real intentions of the program in question, and 2) they encourage the use of implicit rather than explicit assumptions. I would prefer that the minister who wants to "save souls" keep that objective and assume that church membership is a reasonable estimate thereof, rather than that he restate his objective in terms of maximizing church attendance. In the former case he knows where his assumption is and can be sensitive to its influence on his decisions. Reference No. 6 gives some valuable insight on this point.

Alternative Approaches

What are the alternative approaches to accomplish the objectives? There is a developing literature on this subject, much of which will likely be covered in the other position statements. In the Appendix I have specified five attributes which might be measured, depending on the objectives of evaluation: 1) exposure, 2) quality or presentation, 3) change in knowledge, skill, or attitude, 4) behavioral change, and 5) end result change. It may be a moot point but I would like to reserve the term end result to patient health status and to look at behavioral change as a necessary precondition. The reasons for this are twofold: 1) technically patient care is an activity, not an end result; 2) isolated focus on behavior does not shed much light on the priority importance in terms of the over-all health of the target population. Conceivably, extensive efforts might focus on a relatively trivial behavioral set, especially if given the mandate to choose as priorities those activities which are more quantitatively measurable.

Cost Effectiveness of Alternatives

What is the probable effectiveness and cost of each alternative? This question is key to rational decision-making on the economics of evaluation. If one pursues it seriously, it becomes obvious that there is very little relevant data on the cost-effectiveness of evaluation. In fact, faith in the ultimate worth of evaluation is one of the primary justifications as it is in the field of education. Two authors who surveyed the evaluation of health programs in 1968 had this to say after reviewing 221 articles: "A thorough and exhaustive survey of the literature has led the authors to conclude that, first, evaluative research is still in its infancy; and, second, most publications on evaluation of health programs extoll the virtues of evaluative research but offer little operational guidance. The paucity of available research may be imputed to the lack of funds or lack of interest but it is more probably a reflection of the lack of valid and reliable measuring instruments and the formulation of health goals in Utopian language."[7]

Certainly the literature scanned in preparation of this paper was relatively unproductive.

Obviously, any meaningful judgment as to the effectiveness of continuing education evaluation efforts should be tied to the objectives of the study. Existing methodology can and does prove useful in improving course design and educational approaches. And, although much maligned, exposure studies are still extremely useful in justifying the apparent contribution to society even if the focus is on counting input rather than measuring output. Cost is a highly relevant factor in evaluation performed to meet the objectives of improving course design, improving teaching methodology, or justifying program contribution. In these cases the proportional expenditure for analysis should be viewed in the context of over-all program costs with some allowance for serendipity. Cost is less relevant to evaluation done to meet the objective of contributing to general knowledge about educational approaches or measurement techniques. The guiding principle in judging how much should be spent for evaluation is that marginal costs should approach but not exceed the expected marginal value of the information obtained.

The Appendix to this paper was developed to answer the basic question of what kind of costs we are talking about at various levels of evaluation. A one-week short course is outlined at an estimated cost of $42,000. Actual costs would vary with course design, content, and the participant group. It is interesting to note that participant time is by far the most costly item. The following table shows the estimated costs of five different levels of evaluation (details are included in the Appendix).

Estimated Cost of Different Evaluation Approaches For A Hypotehtical One-Week Short Course

Evaluation Approach	Estimated Cost	Fixed Course Cost	Percent of Course Cost
Participant Opinion Analysis	$ 945	$42,000	2%
Pre-post Examination of Knowledge	2,455	42,000	6%
Mail Survey of Performance Change	47,630	42,000	113%
Survey and Observation of Behavior Change	155,850	42,000	371%
End-Result Studies of Patient Health	212,350	42,000	506%

There is a great deal of approximation involved in such estimates, but not much more or less than that involved in developing the actual budget of a research proposal. The participant opinion analysis and pre-post examination are within an acceptable range in proportion with course costs. The studies of behavioral change, or patient health status, greatly exceed course costs and could be justified only as they relate to the objective of contributing to general knowledge.

Selecting An Evaluation Strategy

What evaluation strategy, or mix of strategies, will maximize gain in terms of the value of information received while minimizing cost? Rarely could an evaluation strategy be selected on a purely quantitative basis. But the point of view of attempting to maximize gain while minimizing cost is a useful ideal. There is a danger in this approach of looking only at short-range benefits where the real pay-offs might occur in the long-range future. The final choice should involve a close look at the objectives of evaluation, the likelihood of obtaining useful data, and the estimated costs. There is no general rule as to what cost level is appropriate as compared with over-all program costs. Nick Parlette suggests the general idea of setting aside a 5 to 10 percent override for evaluation, some of which might be accumulated for several intensive research programs aimed at attempting to build evaluation skills and perform end-result studies.[8] The general idea of this approach seems tempting, although the proportion allowed for evaluation may be small given the lack of reported research at this time. However, it would be misleading to suggest that such a percentage be spent for evaluation of each course. It would be fruitless to perform studies which would yield course improvement data for nonrepetitive ventures. It would be equally futile to extensively evaluate a given program when it was apparent in advance that the data would not be utilized. On the other hand, certain courses might lend themselves to rigorous evaluation aimed toward building educational theories. In these the evaluation expenditure could justifiably exceed course costs many times.

Summary

We are currently entering a stage in which highly rationalized, quantitative decision-making will likely guide the resource allocations of most federal, state, and local governments. This can be expected to have a definite effect on continuing education program development which, given our lack of past evaluative efforts, may be of some real benefit. It is important to view this trend with a mixture of optimism and caution, for it could lead to significant improvement in the quality of decision-making, or it could lead toward a rather undesirable state of mechanized mediocrity. Perhaps one of the more noted commentators on the public issues involved is C. E. Lindblom whose article, "The Science of Muddling Through," published in 1959 is still a classic analysis of decision-making in a pluralistic society.[9]

The primary difficulty of analyzing the economics of evaluation in the health field occurs in assessing effectiveness values rather than in estimating costs. One author hints at the difficulty in an interesting attempt to correlate increase in health expenditure with longevity and concludes that: "It is probable that we could either halve or double the money now spent on health without significantly affecting our longevity."[10] His results technically reflect the fact that longevity is no longer a sensitive index in this country (neither, for that matter, are most morbidity or mortality data). The lack of a market-place, the small role of consumer choice, the noncompetitive orientation, and the prevalence of intangible, noneconomic values which resist quantification[11] all lead to extreme difficulties in analyzing the benefits of improved health services. In terms of continuing education, there is even great difficulty in deciding whether costs should be treated as assets (investments) or expenses.[12]

The uncertainties involved in measuring benefits in the health field make it impossible to arrive at a completely objective judgment as to appropriate evaluation costs. Given this situation, it is convenient to borrow a closing remark from a session on the Programming, Planning, and Budgeting System[3] and quote Jean Paul Sartre: "These are difficult times; one does what one can."

References

1. Novick, David (ed.): Program Budgeting. Cambridge: Harvard University Press, 1965.
2. Goldman, T. A. (ed.): Cost-Effectiveness Analysis. Washington, D.C.: F. A. Praeger, 1967.
3. Alfandary-Alexander, M. (ed.): Analysis for Planning Programming, and Budgeting. Washington, D.C.: Washington Operation Research Council, 1968.
4. Mood, A. M.: "On Some Basic Steps in the Application of Systems Analysis to Instruction." Socio-Economic Planning Sciences, 1:1, 1967.
5. Association for Hospital Medical Education. Evaluation of Graduate Physician Manpower Education. Chicago, Illinois, 1969. Seminar and report sponsored by USDHEW Contract PH 108-69-40.
6. James, G.: "Evaluation in Public Health Practice." American Journal of Public Health, 52:7, July, 1962.
7. De Geyndt, W.: Evaluation of Health Programs - An Annotated Bibliography. Minneapolis: System Development Project. Issue 9 of comment series, 1968.
8. Parlette, N.: "Evaluation of Continuing Education." AJPH, 60:2, Feb., 1970.
9. Lindblom, C. E.: "The Science of Muddling Through." Public Administration Review, 19:2, (Spring, 1959).
10. Forbes, W. H.: "Longevity and Medical Costs." New England Journal of Medicine, 227:71-78, July 13, 1967.
11. Somers, H. M. & A. R.: A Program for Research in Health Economics. Washington, D.C.: U.S. Department of Health, Education and Welfare. Health Economics Series No. 7, 1967.
12. Douthat, J.: "Accounting for Personnel Training and Job Development Cost." Training and Development Journal, 24:6 (June, 1970).

Appendix

Cost Estimates of Alternative Evaluation Techniques

Introduction:

This appendix has been developed in order to provide a basis for cost comparison among alternative evaluation techniques. Its purpose is to provide a response to the question of what kind of costs are involved at specified levels of evaluation. Such costs are not fixed amounts subject to precise definition but are reasoned guesses of resources required for specified research approaches.

Individual cost estimates for five research approaches are generated, assuming a hypothetical short course of one week's duration. The resulting data are useful to a discussion of economic allocation issues and priorities. I have attempted to be conservative in the estimates to avoid any inclination to exaggerate such costs. Anyone desiring to use these data for budget estimate purposes is alerted to examine the figures carefully.

Evaluation Alternatives:

There are a number of alternatives for evaluating a short course, depending on the purpose of the evaluation and the related attributes to be measured.

The reasons we might consider for evaluating a short course include:

1. Developmental Evaluation -- To improve course design and presentation.

2. Normative Evaluation -- To compare different instructional approaches in order to choose the most cost/effective educational strategies.

3. Inductive Evaluation -- To contribute to knowledge or build methodology related to continuing education.

4. Absolute Evaluation -- To justify the societal contribution of continuing education as compared with other resource allocations.

The attributes which could be measured and illustrative evaluative approaches include:[31]

1. Exposure (Outreach):

   a. Differential study of participant distribution and characteristics.
   b. Motivation and interest studies.

2. Quality of Presentation (Input):

   a. Faculty judgment.
   b. Participant opinion analysis.

3. Assimilation (Knowledge, Skill, and Attitude Change):

   a. Pre-post examination.
   b. Critical incident studies.

4. Job Performance (Behavioral Change):

   a. Mailed questionnaire.
   b. Field survey.
   c. Professional survey and observation.

5. Societal Result (Output):

   a. Judgmental estimate and analysis.
   b. Direct observation and measurement.

Most reported evaluation efforts have been concerned with differential exposure analysis, studies of the quality of presentation, and examination of participant assimilation. Lately, the developing sophistication of educational theory and research and the increased weight of the management science viewpoint in government decision - making have served to change this emphasis. As pressure increases for greater attention to evaluation, the objectives and comparative cost must be considered in terms of over-all program goals.

The following estimates have been developed given the lack of empirical data on this topic.

Course Costs:

Suppose we were concerned with choosing an appropriate evaluation approach for a one-week (five-day) short course for fifty physicians. The direct and indirect course cost might be estimated as follows:

Instructional Costs (Assumed to be $40 per day per student)	$10,000
Participant Expenses ($20 per day per student)	5,000
Participant Travel (Assume Average Distance 400 miles)	2,000
Replacement Salary or Income Loss (Assume Average Annual Income of $24,000)	25,000
Total Course Cost	$42,000

The actual cost of a given course would of course vary widely depending on duration, content, type of participants, location, and material involved. The above estimate should be reasonably representative of the cost of the general short course in the medical field.

Evaluation Costs:

Given the four purposes of evaluation (course improvement, knowledge change, behavior change, and end-result change) we might select the following evaluation approaches at specified cost levels.

1. Participant Opinion Analysis (Quality of Presentation):* Participants are asked to record their interest in each presentation, estimate its value to their performance, and make general suggestions as to course design and content. The results will have relevance to redesign of the course if repeated and to general course planning. A number of reactionnaires designed to collect this type data are available, thus development is largely a matter of adaptation. 32

Professional Design (1 executive day)	$ 80
Participant Cost (3 man days total)	330
Analysis (5 math aide days)	175
Report Preparation (2 executive days)	160
Support Personnel, Materials, Distribution	200
	$945

2. Pre-Post Knowledge Examination: Participants would be asked to complete an objective content examination at the beginning and upon completion of the course. It is assumed that construction of the test would involve two fifty-question instruments of equal difficulty. The examination would be pretested and refined including verification of results prior to use. It would be administered to a small control group in order to see if the examination itself stimulated learning. The results of any given pre-post knowledge examination are difficult to interpret except in terms of estimated relationship to given objectives. However, the results can have high normative value when data from several courses or educational approaches are used for comparative purposes.

Professional Design (5 M.D. Consultant days)	$ 550
Test and Measurement Design (2 Psychologist days)	140
Pretesting, Refinement, and Development (includes control and refinement)	500
Administration and Analysis (5 Math Aide days)	175
Participant Time (6 days)	660
Statistical Analysis (3 Statistician days)	210
Report Preparation (2 M.D. Consultant days)	220
Support Personnel and Report Preparation	300
	$2,755

3. Mailed Questionnaire Study of Performance Change: Six months after attending the course, participants are asked to: 1) give their opinion of performance change; and 2) respond to diagnostic and treatment alternatives related to critical patient care incidents. The latter would include situations covered in course content and situations omitted from course content. Each respondent would be matched with a non-participant for age and type of practice for control purposes. Questionnaires would be sent to the fifty participants and an equal number in the control group. Adequate and inadequate response patterns would be developed by an expert panel of medical consultants and the test instrument thoroughly pretested prior to its administration. It is assumed that agreement to participate in the post-course study would have been a condition of enrollment. Given sufficiently discrete course content, there might be a possibility of obtaining results which would be useful in estimating the degree of success in achieving stated behavioral objectives.

Professional Design (30 M.D. Consultant days)	$ 3,300
Experimental Design Team (Statistician, Psychologist, Sociologist, Operations Analyst - 60 days total)	4,200
Project Director (168 M.D. Consultant days)	18,480
Math Aide (168 days)	5,880

* The personnel costs are derived from the following assumed salary levels based on data reported in "National Science Foundation," American Science Manpower, 1968. Washington D.C.: U.S. Government Printing Office, 1969 (Reference 1). Specialist Medical Consultant - $110 per day; Executive Administrator - $80 per day; Statistician - $70 per day; Psychologist - $70 per day; Sociologist - $65 per day; Economist - $65 per day; Operations Analyst - $80 per day; Interviewer - $35 per day; Math Aide - $35 per day; Secretary - $22 per day.

Statistical Analyst (60 man days)	4,200
Participant Cost (12 days)	1,320
Computer Programming	2,000
Computer Operation	800
Secretarial Support (225 man days)	4,950
Materials and Distribution	2,500
	$47,630

4. Professional Survey of Patient Records and Observation: In order to obtain baseline data, prospective participants would be interviewed and patient records studied upon enrollment in the course and again one year after attendance. Identical interviews spaced one year apart would be conducted with a smaller control group (say ten physicians), each of whom would be matched with a reference individual in the participant group.[6] Assuming about 3,500 patient care incidents per participant per year, a total of about 210,000 record entries could be expected per year including the control group.[1] Selecting even one percent of these for study of special clinical entities would require scanning of the entire number of entries and detailed review of 2,100 cases for each of two years (given the design of comparing diagnosis and treatment patterns over the year prior to the course with those in the following years). Information gained from review of records would be supplemented by interview and direct observation in practice situations. This type of study would have to correct for bias among the physician group willing to participate and the nonparticipant in order to generalize about over-all course benefits. The nature of course content would have to lend itself to measurement such as involving discrete and relatively new treatment patterns or diagnostic techniques. The atypical nature of such content would also need to be considered in generalizing about study results. The information gained by such a study would provide direct feedback on achievement of behavioral objectives specified in course design. It would be particularly valuable in the inductive sense of contributing to insight on evaluative research methodology or general knowledge about effectiveness of continuing education.

Experimental Design (90 Senior Analyst days)	$ 7,200
Professional Criteria (40 M.D. Consultant days)	4,400
Project Director (2 years)	40,000
Educational Psychologist (18 months)	27,000
Record Analysis (200 man days)*	7,000
Observer-Interviewer (90 M.D. days)*	9,900
Statistician (100 days)	7,000
Participant Cost (1 day each, 50 days)	5,500
Secretaries (3 years)	14,850
Math Aides (2 years)	16,000
Computer Programming	8,000
Computer Operation	1,000
Transportation	5,000
Materials and Distribution	3,000
	$155,850

5. Direct Measurement of Societal Result: For purposes of cost estimation, it is assumed that an end-result study of patient health status would be preceded by analysis of participant behavioral change and that thereafter it would have many of the characteristics of the more familiar drug study. Altered physician behavior would be handled as a treatment pattern (e.g., experimental drug) and evaluated in much the same way. Such an approach would only be practical for a specific clinical entity involving discrete behavioral patterns and measurable change in patient health status (as related co the clinical entity in question). A 10 percent sample would be drawn from the

* The estimates allow the record analysis scanning time of 100 records per hour and review time of one-half hour each for 2,100 records. A physician would be required to conduct interviews and observe practice patterns with the assumption that an average of two interview-observation events could take place each day. Travel time is added.

patient records identified in the study of behavioral change, including a matched control group, for a total of 420 persons. Patients in the sample would be interviewed by a physician who was not aware of whether the subject was a member of the control group or study group and an estimate made of his health state, as related to the given condition. This assumes that interview and observation would be sufficient in this case. The potential value of the information from a thorough study of this sort would be extremely high, assuming that a course could be found which met the necessary specifications. Conceptually, such a study could contribute to an understanding of the nature of the association between educational activity and end result, could measure the validity of lesser evaluation techniques, and could provide some basis for computation of educational benefits. In reality it might be difficult to find a situation in which the study could be conducted.

Study of Behavioral Change	$150,350
Additional Design Efforts (30 days)	2,400
Added Project Staff Time (1/2 year)	22,000
Participant Cost (1 day each, 50 days)	5,500
Patient Interviewers (210 M.D. days)	23,100
Added Analytical Time	5,000
Transportation	4,000
	$212,350

Cost Comparisons:

The cost implications of the alternative evaluation approaches appear when compared with over-all course costs and possible utility factors. If the figures are anywhere near correct, evaluation costs could be expected to run from about 2 percent to 500 percent of the total course costs. It is obvious that the behavioral studies and end-result studies can only be justified in terms of the potential contribution to knowledge about educational design or development of evaluation methodology. The costs of doing these studies for course refinement, course justification, or comparison of alternative approaches is far too excessive.

Evaluation Approach	Total Estimated Cost ($)	Fixed Course Cost	Percent of Course Cost	Potential Utility
Opinion Analysis	945	42,000	2%	Developmental
Pre-Post Test	2,455	42,000	6%	Developmental, Normative
Mailing Survey Performance	47,630	42,000	113%	Inductive
Behavior Observation	155,850	42,000	371%	Inductive, Absolute
End-Result Studies	212,350	42,000	506%	Inductive, Absolute

It is important to note that course costs have been arbitrarily held constant in this example. In reality, these costs could vary widely, while it is reasonable to expect the cost of the described evaluation efforts to remain relatively stable. There would be much less variation in the cost of the behavioral observation approach for a one-day course versus a three-week course than there would in the direct course costs.

The actual decision as to the level of evaluation most appropriate to a given course would involve a number of considerations: the purpose of the evaluation, the probability of being able to measure desired output, the time allowed, the receptivity of the participant, the availability of measurement instruments, and the over-all cost as compared with course cost and expected value of the information. This last factor is undoubtedly the most important and, like many other aspects of education, does not readily lend itself to quantitative treatment. To the extent that evaluation has the purpose of describing the contribution of continuing education to society, each general approach has specific strengths and limitations. Documentation of participant exposure is inexpensive but can be related to a change in societal health state only by an uncertain chain of conjecture and assumptions. Estimation of behavioral change through field observation of participants would be extremely expensive and subject to numerous measurement errors, but it would greatly reduce the assumptions involved.

In the final analysis one cannot escape a basic dilemma common to the human service persuasions in our mechanized culture. On one hand, society is asked to finance extensive educational efforts with very little documentation as to their value. On the other, if it tries to correct this by research and analysis, it ultimately must face underwriting extensive studies with very little documentation as to the value of the information obtained. The obvious recourse in the face of these uncertainties is for society to give greater attention to the tangible and more measurable options. The only practical response is to attempt to develop better estimates of program effectiveness, to argue the intrinsic value of the more qualitative and less measurable aspects of human endeavor, and to bring the best minds to bear on questioning the apparent but deceptive value of the more tangible and visible alternatives. If there is any success, high quality human service might at some time attain the political and social appeal now reserved for our massive highway systems, space programs, defense capabilities, and impressive building program.

References

The following references were drawn upon in developing the cost and physician caseload data used in estimates outlined in this appendix.

1. National Science Foundation. American Science Manpower, 1968. Washington, D.C.: U.S. Government Printing Office, 1969.
2. U.S. Department of Health, Education and Welfare, National Center for Health Statistics. Health Resources Statistics, 1968. Washington, D.C.: U.S. Government Printing Office, 1969.
3. U.S. Department of H.E.W. Health Manpower, Perspective, 1967. Washington, D.C.: U.S. Government Printing Office, 1967.
4. U.S.D.H.E.W., Resources for Medical Research. Washington, D.C.: U.S.G.P.O., 1963.
5. National Advisory Commission on Health Manpower, Report Volume I. Washington, D.C.: U.S.G.P.O., 1967.
6. U.S.D.H.E.W., National Center for Health Statistics. Physicians Visits, Series 10: Number 49. Washington, D.C.: U.S.G.P.O., 1968.
7. H. H. Avnet: Physician Service Patterns and Illness Rates. New York: Group Health Insurance, Inc., 1967.

MEASUREMENT PROBLEMS

Hugh M. Petersen, Ph.D.

The major problem associated with the evaluation of continuing medical education is that of establishing the reliability and validity of measurements. Most of us know something about the statistical definitions of these two terms. However, I'm sure there are those who have not realized that we have initially a problem in semantics. That is, in a real sense we have a problem of knowing the true meaning of our measurements as well as that represented by their statistical reliability and validity.

This problem may be exemplified by considering the reliable and valid measurement of "learning." Most of us would be willing to hypothesize at a theoretical level that "learning" by practicing physicians, through continuing educational opportunities, should improve patient care. We might go beyond the hypothesis in the direction of tacit assumption. In either case, we are operating and communicating at the factual-meaning, or somewhat abstract, level.

At this level the meaning varies depending mostly upon the level of abstraction employed. Many of us would agree as to some of the activities which serve as illustrations of learning; e.g., acquiring a medical vocabulary, memorizing the parts of human anatomy, and acquiring surgical skills. There are other activities which are more subtle examples of learning, such as the acquisition of medical prejudices and preferences, as well as other social attitudes and ideals. Perhaps we should be satisfied to define learning as "improvement with practice," or "profiting from experience." Yet, we know quite well that some learning is certainly not improvement, and other learning is not desirable in its consequences. Maybe a better approach is to define learning as a "change in performance resulting from some specified educational experience." None of these definitions is really acceptable because the level of abstraction employed introduces too much ambiguity. We really don't know what these definitions mean until we reduce the ambiguity inherent in the various terms utilized in their construction.

At the factual-meaning level our communications are fraught with ambiguities; i.e., we don't really know what is meant by "learning," by "continuing educational opportunities," by "improve," or by "patient care." If each of us were asked to define these terms, much variation would result from one person to the next. This fact of ambiguity directs us to look for a more concrete, "point-at-able" level of communication which insures considerable invariance in meaning.

Educators and students often use concepts that are formulated at rather high levels of abstraction. These are quite different from those utilized by the empirically minded educational researcher who operates at the "point-at-able" level. In attempting to overcome the differences between these two levels, it is suggested that we refer to the latter kinds of variables as indicators of the former concepts. In so doing we adopt the position of operationism.

The philosophy of operationism and the idea of defining the abstract in operational terms is commonly associated with the physicist, Percy Bridgman, writing in his work, The Logic of Modern Physics.[1] Arthur Eddington, the astrophysicist and philosopher of science, gives us some insight into the need for operational definitions by means of the following problem. "An elephant slides down a grassy hillside....." The mental picture is probably stimulating, but that is irrelevant. Reading on, we find that the mass of the elephant is two tons. At this point we may ask what is, or are, two tons? Do the two tons belong to the elephant? If we think of two tons as property of the elephant -- confusion of meaning results. Historically, physics textbooks have defined mass as a "quantity of matter." However, it was only opinion that such a quantity of matter might be represented by some accepted measurement technique. Present-day physics students would not bother with opinion. They know that mass is inferred from "point-at-ables" such as meter or pointer readings.

All other variables inherent in the elephant's hillside slide, e.g., coefficient of friction, hill slope, and descent time, may be defined by pointer readings. The lesson to be learned is that, even if we have definite ideas and beliefs regarding "real" objects existing in the external world, these mental representations are useless except as handled by the science of operationism. We cannot handle the measurement problem until the concepts or constructs are replaced by specific operations.

George Lundberg, who championed the operationalist position in sociology, strongly defended the premise that all variables are measurable and that one should not be concerned with hypothetical entities. He further claims that it is erroneous to assume that measurement is not a way of defining things; i.e., it is erroneous to assume that measurement is a process which can be carried out only after the "thing" to be measured has been defined.

Eddington stresses that variables should be defined according to the way they are recognized, and Bridgman points out that the proper definition of a concept is not in terms of properties possessed, but in terms of actual operations. Therein lies the essence of operationism. It means that, if variables are defined in terms of properties, there is no possible way of testing directly any hypotheses in which the variables appear. Tests and experiments must be performed in terms of operations; it is the "point-at-ables," i.e., the pointer readings, that are related.

The essence of operationism consists in reducing a situation to elements with which we are so familiar and unambiguous that we accept them as a matter of course, so that our curiosity rests and overconfusion is reduced.

The use of operational definitions in describing the variables associated with learning insures reliability in the sense that others may duplicate our system of measurement. Thus, in measuring learning, we must devise some meter, representing an objective scale, which we monitor before and after an educational experience. In order to determine whether learning has occurred, we would compare pointer readings as represented by pre- and post-test scores. If a change in pointer readings results, we subject the change to a statistical test in order to assess the probability that chance was the major determiner of change. If the probability of chance being the source of change is small, we conclude that the change is due to learning.

Summarizing to this point we see that concepts and constructs are useless in educational measurement unless they are operationally defined. The act of operationally defining insures the realiability and replicability of our measures.

The positive relationship between operationism and measurement reliability does not necessarily hold for measurement validity. For example, we may operationally define learning as a statistically significant positive change in test scores from pre- to posttest experiences. If our test items have been well designed, chances are they are statistically reliable. Since the test has been unambiguously specified, logical reliability as well as replicability has been assured. The validity of our measure of learning is something else.

More and more we are beginning to accept the fact that there is no real validity in academic measures of learning in medicine, except as may be demonstrated within the context of patient care. The only valid system of measuring learning involves the measurement of the predisposition to apply course content in job performance. We must be able to monitor physician-patient interaction and the results of such interaction before we know anything of substance regarding the effects of medical education experiences.

Let us consider the case of postgraduate courses in hematology. The University of Kansas has prepared a course of study in hematology[3] which is well known and respected by the many physicians who have used it. Much work has gone into the preparation of the course, and yet that work cannot be properly evaluated by restricting our evaluation to the monitoring of pre- and posttest measures of the usual paper and pencil type. This type of evaluation does not do justice to the course or the physician-student.

It is suggested that physician job-performance measures be designed for each continuing education experience. These measures would indicate the extent to which the physician applies course material he has studied in any educational program.

As a specific example of the measurement of physician application of principles and techniques learned in a course in hematology, the following evaluation instrument is presented.

Hematology Course Evaluation

(Parameters to be recorded)

A. General Information

1. Attending physician_____
   A. G.P. _____
   b. Int. _____
   c. Other _____ Specify_____
2. Case number_____
3. Race_____
4. Sex: Male_____ Female_____
5. Age: <10____ 10-20____ 21-40____ 41-60____>60 ____
6. Consultant use: Yes_____ No_____
7. Admission date: Day_____ Month_____ Year_____
8. Discharge date: Day_____ Month_____ Year_____
9. Number of hospital days:_____
10. Death: Yes_____ No_____
11. Admitting Dx_____
    _____
    _____
12. Discharge Dx_____
    _____
    _____

B. History:

1. Complete? Yes____ No____

2. Bleeding recorded? Yes____ No____ Not mentioned ____
   a. If yes:
      1) Source recorded? Yes____ No____
      2) Amount recorded? Yes____ No____
      3) Color recorded? Yes____ No____
   b. If patient is female:
      1) Menstrual history recorded? Yes____ No____

3. Hemorrhage into skin recorded? Yes____ No____
   a. If yes:
      1) Extent recorded? Yes____ No____
      2) Duration recorded? Yes____ No____

4. Dietary history recorded? Yes____ No____
   a. If yes:
      1) Type recorded? Yes____ No____
      2) Amount recorded? Yes____ No____

5. History of drugs recorded? Yes____ No____
   a. If yes:
      1) Type recorded? Yes____ No____
      2) Amount recorded? Yes____ No____

6. Jaundice recorded? Yes____ No____

7. Family history of anemia recorded? Yes____ No____

8. Family history of bleeding recorded? Yes____ No____

9. History of weight change noted? Yes____ No____

C. Physical Examination:

1. Presence or absence of pallor noted? Yes____ No____

2. Presence or absence of skin hemorrhages noted? Yes____ No____

3. Presence or absence of active bleeding present? Yes____ No____

4. Blood pressure recorded? Yes____ No____

5. Pulse recorded? Yes____ No____

6. Presence or absence of "spider" angiomata? Yes____ No____

7. Tongue described? Yes____ No____

8. Liver size noted? Yes____ No____

9. Spleen size noted? Yes____ No____

10. Lymph nodes size noted? Yes____ No____

11. Edema (swelling) of legs and feet noted? Yes____ No____

12 Neurological examination done? Yes____ No____

D. Hematological Consultant:

1. Was one used? Yes____ No____

2. Consultant's diagnosis?
   a. Does it conform to that of admitting physician? Yes____ No____

E. Laboratory Orders:

1. CBCs (or fractions thereof). How many ordered? _____

2. Was fecal occult blood test ordered? Yes____ No____

3. Was there an order for fecal examination for parasites? Yes____ No____

4. Was total serum iron ordered? Yes____ No____

5. Was iron-binding capacity ordered? Yes____ No____

6. Was Coombs' test ordered? Yes____ No____

7. Was haptoglobin ordered? Yes____ No____

8. Was reticulocyte count ordered? Yes____ No____ If yes, how many?_____

9. Was hemoglobin electrophoresis ordered? Yes____ No____

10. Was serum protein electrophoresis ordered? Yes____ No____

11. Was gastric analysis ordered? Yes____ No____

12. Was serum folic acid ordered? Yes____ No____

13. Was serum B-12 assay ordered? Yes____ No____

14. Was Schilling test ordered? Yes____ No____

15. Was red-cell osmotic fragility test ordered? Yes____ No____

16. Was glucose-6 phosphate dehydrogenase activity ordered? Yes____ No____

17. Was pyruvic kinase activity ordered? Yes____ No____

18. Was bone marrow study ordered? Yes____ No____

19. Was liver biopsy done? Yes____ No____

20. Was lymph node biopsy done? Yes____ No____

21. Was initial screening procedure for hemorrhagic diathesis?
    a. PT: Yes____ No____
    b. PTT: Yes____ No____
    c. Bleeding time: Yes____ No____
    d. Clotting time: Yes____ No____
    e. Platelet count? Yes____ No____
    f. Thrombin time? Yes____ No____
    g. Clotting time? Yes____ No____
    h. Clot retraction? Yes____ No____
    i. Test for fibrinolysins? Yes____ No____
    j. Prothrombin consumption time ordered? Yes____ No____
    k. Was fibrinogen level ordered? Yes____ No____
    l. Was chromosome culture ordered? Yes____ No____
    m. Was leukocytes alkaline phosphatase ordered? Yes____ No____

This hematology evaluation form probably needs some refining, but it represents a concrete attempt at operationally defining the learning which may have taken place as the physician has studied. The comparison of pre- and postcourse physician behavior would yield some unambiguous pointer readings, representing validity far beyond that inherent in paper and pencil tests.

References

1. Bridgman, R. Q.: The Logic of Modern Physics. New York: Macmillan, 1928.

2. Stevens, S. S.: In Theories in Contemporary Psychology. Melvin Marx, ed. New York: Macmillan, 1963.

3. Wilson, S. J., and Larsen, W. E.: Hematology-Basic Course (home study course). Department of Postgraduate Medical Education, University of Kansas School of Medicine.

SECOND PLENARY SESSION

CHAIRMAN RISING: It is the plan, subject to your pleasure, to devote this morning's session to thirty-minute presentations by each of the discussion groups of yesterday. During that time I am going to turn the lectern over to the Group Consultants, with the exception of Steve Abrahamson who had to leave, in which case Bill Nelson has kindly consented to be the substitute Group Consultant. The Group Consultant may use his time in any way he sees fit. He may give a crystallization of his group's discussion of yesterday, he may ask somebody else to give it, or he may call on sundry members of the group, even all the group, for their contribution.

This afternoon members of other groups who want to enter arguments, other ideas, modifications, etc. will be free to get into the discussion. So, unless I hear a majority wishing to change that, we will reserve the afternoon for general discussion by everybody here. We hope to have all of the different points of view represented by the persons of varying backgrounds and interests reflected both this morning and this afternoon.

Group One will give the first report.

DR. NELSON: It was a challenge to put our group's discussion into some sort of order. We covered a lot of territory. What I would like to do is put it under several headings and give what I think was said in each of several different topic areas.

The topic of the day was "Obstacles to Continuing Education" and our particular group was supposed to be concerned about "Inappropriate or Unrealistic Goals." As you recall, Steve presented some of those to us yesterday morning.

The net result was probably a discussion, if I may paraphrase it, of various problems in pursuing continuing education programs, and perhaps "problems" is the same as "obstacles" but I am not sure it is exactly the same.

The comments fell into five main areas:

1. We talked about motivation, or lack thereof, to pursue continuing education.

2. We talked about the problem of wants versus needs -- that old saw that we have been down the road a good many times on -- the wants versus the needs for continuing education, and both the recognized and the unrecognized needs.

3. We talked briefly about the pressures that are developing toward evaluating wants or needs, not at length but at least we recognized there were some pressures in this direction, and I think it is very important.

4. We talked about techniques for evaluating needs. I want to point out we did not discuss evaluating the educational programs themselves.

5. Very briefly there were some comments on techniques of delivering continuing education, of meeting the needs. It was brief but it focused on some interesting things in the national picture.

So, if I may, I would like to report on these five categories. I am sure members of the group will have comments to make in amplification of some of the points that we discussed.

Motivation was the area in which we started. Someone asked the original question, "Isn't motivation the main problem?" Then they or someone else went on to indicate that the educational opportunities are there for continuing education for physicians, in particular, but asked whether it isn't a question of motivation whether physicians take advantage of these opportunities. Indeed, don't we see, really, that physicians may pursue continuing education in medicine, but they also may pursue golf or they may pursue stocks and bonds, as somebody put it, or many other things. Isn't motivation, then, the question?

This led to a discussion of the problems of priorities on the part of physicians for their time. They are concerned with these other things, but they also are concerned with delivering good quality care too. There is a variation in this, but the physician is faced with a balance in terms of the many demands on his time, the "busyness" of his practice, the need to have relaxation versus the question of pursuing it in terms of continuing education.

Steve made a very strong point that he felt most physicians do want to practice good medicine. I would like to emphasize that I felt strongly this was true. I believe most physicians do want to practice good medicine if it is possible, and they are motivated to want to improve. However, they are faced with this question of priorities and to what degree they want to put the effort into the continuing education efforts versus perhaps other things.

Then Steve went on to indicate that part of the problem he felt was inappropriate objectives in education and in the things that we offer that perhaps kill the interest (motivation) of physicians to be involved. A little bit later we asked questions about this, and I think what he really was saying is that many of the programs that are offered really do not meet the needs of the physician as the physician sees his own need.

Someone brought up a list of interfering factors with regard to motivation. I will not recite them all, I am sure most of us know them, but they are the kinds of things that were in the Vollan Report[1] of many years ago and the California study, and various other studies. They concerned the problems of the "busyness" of his practice, his inability to get away, the finances, the timing, the availability of programs for what he was interested in, etc. There are many of these factors.

Another factor was the learning process. The study process itself is basically toil, it is work, it is effort, and to be involved in continuing education is to face up to doing something that is not necessarily a whole lot of fun. This raised the question that perhaps one of the ways to overcome some of this difficulty is to be sure that we think in terms of programs that are attractive.

Finally, under motivation, one of the members noted that medical students under certain situations tend to shun taking advantage of the learning-study opportunities which they have but which are not required in their course efforts, and this led to the suggestion that perhaps in some ways we are tending to breed a group of physicians today and in the future who will be even less motivated than physicians in the past.

Next we discussed wants and needs. In this category it was noted that physicians vary widely in their practice activities. This is stating the obvious, but it is important and there were a lot of different factors brought in. "General Practitioner" is a very broad term and there are many kinds -- the general practitioner versus the internist, rural versus urban -- I don't think it was put in those words but I think that is what it meant.

1. Vollan, Douglas D.: Postgraduate Medical Education in the United States. A.M.A., 1955.

One very interesting point that came up was that there are differences in what physicians need to know in rural areas versus urban areas. This came up, incidentally, in connection with medical audit that we will comment on in a few minutes. Maybe there are differences here. One of the spokesmen who spoke effectively from the standpoint of rural practice said, "I do not believe that is significantly true. Today, with communications and exchange of opportunities, with the mail for steroid determinations, etc., there isn't much difference between rural practice and urban practice in terms of what we need to know."

The second "wants and needs" area of discussion was needs that we now tend to observe in analysis of hospital records. Indeed, are we not missing the boat because office records are very different, and yet ninety percent of all practice is office practice. I think it might be pointed out that most medical groups are concerned with office practice. The American Society of Internal Medicine, as many of you know, is in the process of making quite a significant study in Syracuse.

Wants and needs are different, and Bob Neth made a couple of brief comments on this. I will not try to paraphrase what he said except that he pointed up the difficulties that may be involved in replying on the wants and the popularity expressions of individuals with regard to some of the course programs in which they have been involved.

Steve Abrahamson indicated that our thoughts regarding the physician's wants and/or needs may be quite different from his, and this may lead to inappropriate objectives in our programming.

There was a brief discussion regarding the role of basic sciences in terms of wants and needs. I will leave that for others to talk about. There was somewhat of a difference of opinion about the necessity of being unduly involved in basic science.

It was pointed out in regard to wants and needs that just the plain ordinary routine history and physical, and many other routine things need to be given more consideration than they sometimes get. In this regard certain individuals recalled the role-playing approach that the Board of Orthopedics is taking in terms of evaluating the capacity to do history and physical effectively.

In addition to these wants and needs there were two or three that were very interesting because they were a little bit off-beat and maybe we do not always think of them.

In terms of this total list, somebody said, "You know, we have been discussing the medical tangibles, pneumonia, the capacity to handle the acute coronary, whatever you will. What about insights into the emotional areas? Not just psychiatry but emotion in relation to illness, in relation to life and death, in relation to conception, to family business, etc., the kind of things that we do not really see even in the records. These are the things that go on between a physician and his patient, or the patient's relatives. Do we not need something in evaluating procedure here?

It was mentioned just in passing that there are a number of courses for practicing physicians in relation to emotions, not necessarily psychiatric, and they are often well attended. One of the important ones that I know is Dr. Kauffman's at Mount Sinai. We have them in Albany and I suspect many of you do too. In that regard it was asked, "How do you visualize a medical audit in this area?"

Someone said that often the biggest gap is not in clinical knowledge but in office management. Physicians are being involved in a new role as executives, as one of the individuals put it, in the health care delivery system. How do they answer the phone? How do they use their office personnel? How do they get more efficiency out of everything that they do? Well, it is off-beat, but at least I leave it with you as another area that we need to be a little concerned.

We went on to the question about the wants and needs of the physician with regard to how they use office personnel to help them in educating their patients. This brought up the question of the use of ancillary aids as well as audiovisuals. As many of you know, in pediatrics, obstetrics and gynecology, in maternity care, etc., these are being very attractively and significantly used by some individuals, but there is a vast lack of knowledge on the part of the practicing physician, probably on the whole, in terms of this area. 45

Another one was that many physicians do not know what is available in the variety of allied health personnel services. We all recognize there are a lot of special kinds of services that physicians often do not know are available to them.

Pressure toward evaluating needs was the next topic. There were comments on the fact that there are real pressures building up. These are the third-party payors. (unions, Blue Cross, as well as other public altruistic groups) that are simply concerned with quality of care quite independently of the finances, etc.

This area led to comments upon the possible distinction between recertification and relicensure. There is an important difference. Recertification labels the individual as a person practicing a certain kind of quality of practice. It is not an essential credential for practice. It is not a legal instrument. It is one of quoting the quality.

It was suggested that maybe recertification with challenge examinations is appropriate.

In licensure, because it is a legal matter -- a legal capacity to earn a living -- there is pressure in a different direction, not for challenge examinations to relicense but for requirements that individuals be involved in continuing education experience. If they at least participate they can be relicensed.

Next is the matter of techniques for evaluating needs. Bernard Dryer's "Lifetime Learning for Physicians"[2] was brought up, and Steve Abrahamson had a very interesting comment to make. He said, "Maybe one of the problems with the university without walls was that it was self-instruction without self-evaluation." It is an interesting point.

Steve went on to suggest that self-assessment means regarding patient care, not just self-assessment with regard to knowledge. He was concerned with patient care.

This brought up the question of hospital audit and how to define the pattern of quality of care by which a clinical entity at a given hospital will be defined. I know Clem Brown feels rather strongly that it is important for the hospital staff to be involved in defining the criteria for defining quality care. It is important that the staff be involved in it partly because that is part of the education.

With regard to techniques for meeting the needs, there were brief comments that the instructor should be able to identify with the community, and it was pointed out many instructors do not do this.

There was a question about the status of the instructor. How important is the senior instructor versus the young man who knows his onions?

Also, there was a very interesting discussion regarding developing audiovisual capabilities. I want to point out that Pittsburgh is using audiovisual instruments in the Ob-Gyn instruction program.

I would like to have others from our group comment.

2. Dryer, Bernard: Lifetime Learning for Physicians. Joint Study Committee in Continuing Medical Education. 1962.

DR. CHEZ: I would just like to compliment you on the job you have done. I am totally in awe of anybody who can summarize that much conversation.

MR. HANSEN: I think you did a good job.

CHAIRMAN RISING: Quit complimenting him and argue with him. You know, Bill, if they really agree with you this much I marvel at the fantastic power you have over them!

Thank you very much, Bill.

The Consultant for Group Two was Clem Brown.

DR. CLEMENT BROWN: I will turn our report over to Paul Cudmore.

DR. CUDMORE: Thank you very much, Clem.

Half-way through yesterday afternoon Clem announced that he was not making the report. Somehow I find myself here -- wondering why.

I do want to give you our conversation under three headings. Clem's position paper was on Educator Obstacles, or Educator Problems, but also I am going to talk briefly about the practitioner problems, and then I am going to make a comment or two on institutional problems. I will start with the one that was our prime responsibility, the educator.

I am going to put a few words on the blackboard because these seemed to be important words:

A. EDUCATOR OBSTACLES

Priorities. Really I think our group challenged the group and challenges you. We are, as you know, in the business -- how much are we doing? We are saying quite a lot. Personal priorities of educators was our top thing.

Then, uncertainity about results. Here, supposing we evaluate, supposing we improve our educational system, what evidence do we have that this, in fact, changes the longevity or the well-being of the population? Our group had some real doubts about this. Is there evidence that will help us resolve some of our own uncertainities of the results of our process in any way?

Funding: As educators we have a measure of control over finances. How much have we devoted to this aspect?

There is another aspect to funding and this, again, is one of the prime concerns of our group. If I have a program and I have brought in staff and I have invested a lot of my own personal time and effort, there is a bit of reluctance, there is an obstacle in me to do soul-searching evaluation which may, in fact, not justify the funding, or not justify my expenditure of time. I think we have to be honest about this and say that all of us are probably subject to this negative factor or this obstacle to doing honest evaluation with that vague uncertainty within ourselves that our funding may collapse, our program may collapse, and all the people we have brought in to work with us may be essentially unemployed.

Responsibilities. What the group got at here, was that the person with the responsibility, the educator who has the responsibility for evaluation, may have little influence on the outcome of the program. In other words, he is not the director, he has not the authority to change. For those in our group who are evaluators but not program directors, there may be a communications problem around this.

Money. I think I may have mixed these two a minute ago. I am now referring to money, the fact that we have not allocated enough of our own budget for evaluation, for this obviously points at educators who do not have the attitude that we talk about.

Models. Educators have an obstacle in lack of a good model. There are not good models to which those of us who are in the field can quickly turn for help. Our group felt this, again, was one of the priority items, and that it would help all of us if there were real good evaluation models and these models were readily available; in other words, the next thing is the literature. Another obstacle to the educator is that there is not helpful literature, so someone must do some models and write it up.

Skills. The last point that I wish to make here is the lack of skills on the part of ourselves as educators. Some of us lack the skill to conduct competent evaluations. I think others in the group felt this was not such a big factor. I do not know, maybe we hide behind that one.

That pretty well finishes the list of educator obstacles.

B. PRACTITIONER OBSTACLES

Attitudes of Practitioners. In spite of, or with the full knowledge of Neal Vanselow's paper yesterday supporting a lot of positive attitudes on the part of practitioners, we felt there was evidence of a negative attitude on the part of practitioners toward evaluation. We felt this stemmed from medical school, and all the other schools that we have attended where evaluation is considered almost a weapon of the faculty and an obstacle to those of us trying to get ahead. Somehow, on this concept of evaluation, when you talk to a physician you immediately stir up anxiety and a bit of a negative attitude toward the whole process. We have to deal with this, as I say, in spite of the evidence from the Michigan survey that said the fellows were in favor of evaluation, or at least it can be implied from that. It is one thing to say you are in favor of it, but you know we say we are in favor, as educators, of a lot of things in evaluation and yet there are a lot of things here on which we are not taking any action. To say that the physician is in favor of evaluation is one thing, but to actually have him involved in it is another. I had a personal experience here and I brought this up to the group yesterday; we have found a degree of reluctance on the part of physicians to enter into really soul-searching assessment of what they are doing.

Educational Needs. There is too little information getting through to physicians. We may be talking about this, but there is not very much information getting through to physicians and others that suggests or supports the concept that an evaluation of their educational needs is the beginning of an educational process which is efficient, and that the efficiency of the educational process really demands at least the beginning of an assessment of needs. We are not selling this concept very well. If we are going to get at the attitudes and get to the practitioner, we have to promote this concept a little more. Until that time physicians will be a bit reluctant.

C. INSTITUTIONAL PROBLEMS

Lack of Money. Institutionally there is a lack of money, our medical schools, our various organizations across the country, federal funding and everything, there is just not enough money being focused on continuing education to develop the models and to do the sort of things that have to be done with evaluation.

The System. I mentioned this briefly under educator problems, but I have to come back to it under the institutional problems. Our system is such that even when a problem is identified, the system does not quickly incorporate corrective action, or begin corrective action.

We heard yesterday of the American Heart Association's program on heart sounds. This is just one simple example of this very institutional problem or obstacle in the present system within institutions. The same course continues to be given the same way it was given in spite of the evidence from competent evaluation that it does not work. You know there is something in our whole institutional organization that does not take note of the results that we produce when we do evaluate.

Someone mentioned in the meeting yesterday that there was no over-all institutional strategy. I believe that sums up this. There isn't any strategy in most of our institutions to implement change when evaluation suggests a change should be made.

Geography. Patient care evaluation, which is the kind of thing that we all hope to see happen -- although some of us recognize that intermediate steps may have some relevance -- is a geographic problem for those medical schools and other programs that have a national audience and also a type of local teaching organization. They are set up in Kansas in one area and it is extremely difficult to do patient care research across the length and breadth of this nation, and I would like to include Canada in that, as I happen to come from there. But the geography of evaluation is a tough institutional problem at which we must look.

Medical Records. The phrase that was brought out in the group was the non-system of medical records. If we are to do patient research, then someone has to develop a system. There have been attempts. In most of our institutions -- medical schools, hospitals, offices -- you would really have to buy the deal that there is a "non-system." That is one of the big problems into which we must look.

We have identified the obstacles. We did want to look quickly at some solutions. Call a spade a spade, it is the attitude of the educator. We have to change or this whole thing will not go anywhere. That is the first priority. As educators we must put our money where our mouth is.

The practitioners, their attitudes, here the medical school has a role, and I apologize to those of you who are not physicians. I tend to keep thinking medicine. I hope you can bridge the gap with which I seem to have trouble. We have to introduce or we have to emphasize learning exams. Those of us in continuing education must stand up and shout loudly that the undergraduate medical school bring new emphasis on learning exams and the concept that you learn efficiently when you identify your own weaknesses. This is where it all starts. You set your personal objectives on the basis of an analysis of your weaknessess, and exams then are a helpful positive tool in learning, not a negative hurdle or an obstacle. I hope this can come from this group to all of our schools. Emphasize the positive educational value of learning exams. Create the attitude within the student that evaluation is a tool toward life-long learning.

Institutions that have some expertise in some area must be encouraged to develop that theme beyond where it is now and to let the rest of us know. That is going to require funds, and I hope somewhere, presumably at a high level of federal support, that the kind of money can come that will allow this sort of model to develop within some of the institutions that now have experts in the field.

This has been only the highlights of all we talked over. Clem, I now throw it back to you to conduct the balance of our time.

DR. CLEMENT BROWN: I would like some comments from others in our group.

DR. EISELE: Clem, Paul did a remarkable summary job.

One of the things that concerns me more than other people in our group is the fear of the restrictive effects of the demands for evaluation. I fear the demand for evaluation will force the educational efforts into molds which lend themselves to evaluation but which may not necessarily be relevant or pertinent to actual practice, and we will put the emphasis on some conditions which are not very common that are easy to evaluate rather than on some important things which cannot be so readily evaluated.

Someone mentioned the heart sounds training which in six months evaporated because it was not useful or was not used.

Also, someone mentioned a course teaching people how to diagnose pernicious anemia. This is a rare condition, and certainly relative to iron deficiency anemia, it is very rare. The best way to diagnose the cause of iron deficiency anemia is by doing examinations of stool specimens for blood. This is a very unpopular kind of a test in any hospital I have ever been associated with. The importance of iron deficiency anemia is many times as important as pernicious anemia.

The rigid demand for definitive evaluation may actually restrict the relevant kind of continuing postgraduate education.

DR. THOMAS BROWN: I think I can share Wes's concern that although we definitely do have to develop our models, I do not think we can look exclusively at the cognitive end of evaluation. I do believe that we have to make sure that we build into the evaluating models the affective end of education. But my concern is that we take a total look at evaluation, not just an isolated look.

DR. CLEMENT BROWN: Paul did a fine job. He really did cover most of what we had to say.

Under educator problems, our problems, you listed "skills" at the bottom, and I thought there was more concern with that in the group.

I have some troublesome data, for whatever they're worth, from yesterday that says 40 out of 47 responding said that success of CME programs must be measured in terms of improved patient care, and 38 out of 40 said, "Our medical staff and I would like to achieve the skills necessary to extend our CME program on patient care needs and deficits."

Perhaps you do not really believe that, but it does not seem sensible that someone would say "yes" and not honestly feel there is a problem. I think this is a substantial problem.

The "educationalists" here -- everyone here is an "educator" -- the educationists here know this, but for those of you who are not educationists there is a delightful little book titled, "Developing Attitudes Toward Learning" and its is written by Robert Mager. What I wrote in my paper is somewhat stolen from this. To improve attitudes toward something, and here we are speaking of continuing education, he points out that you have to improve the conditions surrounding evaluation, you have to improve the consequences of evaluation, and you have to have some models of evaluation. That is the subject under question.

As I re-read my paper, and looked at the conditions surrounding the consequences following evaluation that are aversive, that would tend to decrease evaluating activity, there are some very strong ones, pretty unpleasant kinds of things, like you evaluate your course and find out there is no behavioral change. That is a pretty tough one to swallow. Or, if you evaluate your course, and study its effects on the learners, and it comes out that the learners have not changed at all, even by way of a paper and pencil exam or in terms of their quality of care. That is an unpleasant kind of thing.

One of the little deathless quotes that Mager has in here is, "Things surrounded by unpleasantness are seldom surrounded by people."

There are a lot of outcomes of evaluation that are pretty unpleasant. One of the most astounding is: if we evaluate what we are doing we may recognize that our learning experiences are totally inappropriate to the objectives in the first place, and it might suggest to some of us that we need to do something about how people learn, all of which is a shaky kind of concern. There are many good reasons why people do not do any evaluation, and I am sort of skeptical about how much of it is going to be done. However, I do think it is very important.

DR. LEMON: It is interesting that we do not want to bring in physicians and teach them to do something that they are not going to be able to practice. From some of the things you have said, and some of the things you are planning, you apparently are talking in terms of our taking the time to learn skills for evaluation which we are not going to practice until we get some money. I do think we ought to keep that in mind.

CHAIRMAN RISING: Neal Vanselow asked if we wanted to go on to Group Five, since we are on attitudes and his group discussed attitudes. It seems like a reasonable suggestion if we are going to be flexible, and be able to modify ourselves according to the way things are evaluated. Does anyone think that it is not an appropriate move? In this case, then, Group Five will have the stand.

DR. VANSELOW: I am not going to talk very long because I talked long enough yesterday, and we have divided it up so a number of other people in the group will help to present our discussion.

Group Five was given the task of discussing the practicing physician's attitudes toward evaluation. I would like to mention a little about the composition of the group as one thing. We did not have in the group anyone who could represent a practicing physician, so, whatever we say, you might keep that in mind. There were twelve to fourteen people, half of us were M.D.'s and about half were non-M.D.'s, again, no practicing physicians. There were a lot of people from RMP represented in the group, five or six I believe. What particular point of view they would take, I am not sure. I would not know into which category to put them.

We decided that I would make a few introductory remarks, and then ask Frank Lemon and Jim Coole to report on the meat of the discussion.

The first point I would like to make is in rebuttal to Paul Cudmore. If you look at my position paper the point I was trying to make was that physicians are willing to accept the relatively superficial types of evaluation, but if you look at the data again you will see that what most educators consider to be real evaluation, that is patient care evaluation, practicing physicians are not very willing to accept. I think there are many, many concerns practicing physicians have about this level of evaluation.

There are two points I would like to make. The first is the matter of simulation of the patient care situation. This was brought up early in the discussion and pointed out that Steve Abrahamson had Sim One type of model. Also, there have been a lot of paper simulation techniques which have been used recently; the pencil and paper type of thing with erasures and branching patient care problems. The question was, "would this type of technique be more acceptable to the practicing physician than actually going in and looking at his records?" Most people think it would. If there were a correlation between pencil and paper simulation and what the physician actually did in practice, this would be one way around this particular problem.

Someone mentioned there was a very high correlation between pencil and paper simulation and actual audits of records. A few of us were not quite sure of that, and when we get to the discussion later I would be interested if anyone actually has some data.

The other thing I wish to say is that the one thing the group could agree on was that we are having a very difficult time selling meaningful evaluation to practicing physicians. We spent an hour discussing this, then an additional hour proposing some possible solutions to the problem.

I am going to ask Frank Lemon to present the first part of it, and then ask Jim Coole to present the second part.

DR. LEMON: One thing that did not come through but should have is that I feel like I am representing practicing physicians. My background includes general practice in a rural area of Wyoming, and, more recently, the practice of general internal medicine in southern California.

I have been bemused as we've talked, that what is really needed is a controlled study on whether it makes any difference to go to any doctor. Obviously that has implications for continuing education because if it does not make any difference continuing education does not make much difference, and evaluating it would not make any difference at all.

I think our group in general would say that physicians' attitude toward continuing education evaluation is unhealthy.

If physicians are considerably apathetic toward current continuing education efforts, their attitudes toward evaluation of themselves as an essential component for gauging educational needs and for designing programs, are downright suspicious, resistant, and in many instances potentially hostile. That is not a very cozy environment for the continuing education salesman and evaluator. It might even set the scene for another "Death of a Salesman." There is no time for an ineffective soft-sell approach, and a hard one looks as though it might be hazardous.

As Dr. Vanselow has said, we first agreed that continuing education evaluation does not usually sell well, at least the effective forms of it. Then we ask ourselves, Why?

The reasons we came up with certainly cannot be all of them, but we thought strongly that there were two major reasons.

First, the areas in which there was strong agreement could be summed up, again, in one word -- it is a threatening enterprise. Threatening, how? Some saw it on behalf of the doctor as threatening to his own self-image. I spoke about the mornings when I felt engaged in insanity, and that is an uncomfortable morning. No doctor in practice likes to feel uncomfortable in what he is doing that he does not really know the significance of, and is not fully in control of. Evaluation may focus attention on that as a fact of life.

Others saw this threatening situation as his in-built bias against evaluators, somewhat related to what Dr. Cudmore spoke of a few minutes ago. Of course, that might be evaluators in general, it might be evaluators specifically from certain agencies or certain enterprises.

In that connection it was felt there is a rub-off to the continuing education enterprise and to the evaluator in that field of the well-known "town and gown" syndrome. It may very well be that in this area we are getting some of his vote of "no confidence" in the university's perception and understanding of the problems of private practice. They really are two different worlds. I can appreciate this better than many others in the room because I have had my feet solidly planted at different times in both worlds.

Others felt it was threatening because there was a positive fear of the misuse of data so collected by people right in his own community and his own staff, his peers, not to mention the misuse of data that is potential, and for which instances can be cited of the misuse of such data by 'outsiders.' This reflects to some extent a doubt about the confidentiality which we may try to assure and the anonymity which we may promise.

There was a fear expressed by others that he feels this may somehow be involved in the growing process of "the control of medicine." It may be so used.

One could sum up this threatening thing by saying that in an era in which doctors are under attack -- and let us not forget that that is the environment in which we now operate -- it is not a very propitious time for going around auditing what he is doing in the hospital and in his own private office.

Dr. Cudmore brought out what we thought was a second very important area in our group. I'll say it in different words, but it means the same thing. That is, evaluation of the type that we are interested in is not a part of his undergraduate experience. He has not been trained in this as a way of life, self-evaluation for the purpose of self-maturation, self-growth, identification of needs, etc. Undergraduate examinations are feared, detested and, as all of us know, are even inaccurate. Failures in the undergraduate system mark the individual as a failure, but do not mark the system and the evaluators and the educators as failures. Therefore, we come along and talk about evaluation, and the purpose of our evaluation is not understood because it has not been a part of the way of life. Perhaps we are unsuccessful and do an inadequate job of trying to indoctrinate him as to the purpose of our evaluation.

He does not believe that the evaluation is necessary in many cases. This is another reason given. He does not believe that you and I -- particularly those of us from the university who are there because we do not know how to practice medicine, and most particularly those of you who are not educators in medicine at all -- can have any understanding of the art of medicine, and could not possibly evaluate something which will not yield itself to numerical values, things that involve intuition, patient relationships, judgment, the balancing of pros and cons and pluses and minuses. How are you going to measure these things?

Finally, he does not think much of evaluation, probably because there has been inadequate feedback or no feedback in his past experience from that evaluation which could affect his thinking and planning. One more thing, when you are talking to him about evaluation and utilizing his records, he is thinking, "How can my records be compared with anyone else's down the street or over the state?"

There were some miscellaneous reasons, other than these two or three major ones, which were given and which we thought ought to be given more thought.

Evaluation is time-consuming for the physician, and thus annoying. There was disagreement in our group about this, but it was interesting to me that the disagreement seemed to come mostly from people who had not been in medicine and who argued that if you could find time for continuing education you could find time for evaluation. I do not disagree with that, but in this day and age one simply has to realize what a private practitioner's life is, and what his feeling about paper work is in order to appreciate the significance of anything that consumes his time in paper work.

Another miscellaneous reason given was cost. This was looked at from two viewpoints by those of us who considered it. In some instances the fancier evaluation efforts cost to the physician -- in terms of both money and time -- and cost otherwise to the various agencies which might be mounting these kinds of efforts when it is recognized that these do cost, and as a citizen and taxpayer, the physician wonders who is footing the bill.

I suppose it would be only natural, since our own ego status is involved, that we seem to put off until last the reason -- the real reason -- why evaluation is not successful. Maybe we are just poor salesmen. This should have been first, as I have said, but we have put it last.

Dr. Petersen raised the question yesterday very vigorously, having first said that behavior change is the only thing that is worth evaluating. If behavior change is the thing to do, of what are we fearful? Why don't we get busy and do what needs to be done? The answer from our group seems to be that we are fearful of the physician's attitude.

MR. COOLE: It is rather an awesome task to follow a presentation of problems with the presentation of recommendations, especially when our group brought up two rather negative means of assuring evaluation of a relatively stringent form.

The first of these was the suggestion that we need something in the way of mandatory re-examination of physician competency.

Another idea was the task of trying to sell to third-party payors and the public the need for physician competency, and by exerting pressure through these two mechanisms it would force the physician to submit to self-evaluation, and outside evaluation.

There was no real consensus on either of these items within the group, and if anything could be said as a result of the discussions it was that the group felt that some periodic review of a physician's competency was desirable, but, on the question of whether or not this should be mandatory and whether or not this should have any punitive features there was no agreement. There was a relatively heated discussion for quite some time about this one point.

On the positive side, however, I think we did come up with some ideas that bear further examination. The one that I personally feel needs to be promoted the most is the necessity to expose medical students very early in their training to the idea of self-evaluation. It is necessary to promote with these physicians-to-be the attitudes necessary to carry them through a life of self-learning, self-assessment, and self-evaluation, which would make these other negative aspects unnecessary.

In order to do these things we need much more data on physician attitudes about continuing education and the evaluation of this education, and the reasons that lie behind these attitudes. We have a kind of visceral feeling that physicians are antagonistic toward continuing education, or antagonistic toward evaluation and evaluators. They give us this impression. We feel this is true, but we have no solid data to back this up. What is more, we have no data to indicate why these attitudes do exist.

The real core recommendation that the group seemed to feel lay behind all of this was the necessity to involve the practicing physician in the management and evaluation of continuing medical education. We need to sell the idea to the physician that he has to take an active part in both planning and evaluating any program of continuing medical education, be it from a journal club to a formal postgraduate course attended by people from all over the country.

Also, we realized that we need to educate the educator. We need to present to him the aspects in education that would enable him to be a more efficient educator and evaluator. It is an old saw in education that those who cannot teach become principals; and those who cannot be principals go to college and teach education. I think this is becoming true in many fields. Dr. Frank Lemon alluded to the idea that physicians think that those who cannot practice medicine go and teach medicine.

There also seems to be a feeling that if you are an outstanding scholar, if you know the field very well, you are automatically a good teacher. Being an educator, I cannot accept this. It is time we realized that many of our physicians who are in the field of education are good physicians, they know medicine very well, but they are poor teachers. Let us give the skills to be good teachers, and maybe this will change some of the attitudes that physicians have about continuing education and the evaluation of that continuing education.

We also need to convince hospital administrators to accept the responsibility for implementing continuing medical education for their staff, and this implementation needs to be in the form of verbal support and monetary support. We need to convince these administrators to do part of the selling job for us. We can do this by showing them the benefits that can be derived from a solid continuing medical education program.

There needs to be a greater relevancy of evaluation to the patient care that a physician offers.

We have all done the happiness scale, the popularity thing. We got a little deeper in pre- and post-testing. Frequently we have done this and never told the physician what the results were. I think one of the important things in any evaluation of his continuing education is that we tell him what we found out about what he has been doing, but many of us never do this.

We also felt that the area of needs assessment is neglected more frequently than it should be.

Finally, there needs to be some form of in-service training for evaluators to keep them up to date and on their toes about the use of the data that they collect, and the necessity for maintaining relevancy in the educational program in which they are involved.

I trust that any of the rest of the group that has any disagreement with my summarization will at this time let themselves be heard.

DR. VANSELOW: Does anyone else in the group feel moved to comment?

MR. JENKINS: Neal, a comment that you made concerning the pencil and paper tests to replace audits: I think I heard you say that we felt they could do that.

DR. VANSELOW: I think there were several people in the group, or at least one I can remember, who thought there was a high correlation and there were many who thought there was not.

MR. JENKINS: As I recall, Jim Coole was the one who brought this up. There have been a couple of studies but they were fairly selective, there was no real randomization to show any correlation that it can, that there may be a possibility, but there was not adequate proof at this time for us to make that statement, that the paper and pencil test could replace audits.

I am not sure that we came to the conclusion that this is not the time to do audits, because of the conflict that everyone seems to be under the pressure. At least I, as an individual, would feel that despite the pressures we are under from all around, we cannot put these tasks off, that now is as good a time as any. We may never reach the time when auditing would be an acceptable practice, so we have to show how it can do the job and sell it as an acceptable tool, and then go about it. I do not think we can continue to wait just because someone has some anxiety. I think we have to attack the problem.

DR. VANSELOW: Any other comments?

I think one point I might make is that the sense of urgency in this business came up yesterday, the fact that people are considering recertification and relicensure, and the federal government may be getting into this. We do not have a lot of time to waste getting some sort of a system of evaluation of what we are now doing into effect.

CHAIRMAN RISING: Thank you.

All three of the group reports so far have had reference to money: One group called it "funding," another spoke of it as "costs," and the word "money" was never actually used. It was referred to both as a problem and as a solution. It is quite appropriate that we did have those three groups present tandem, and it is even more appropriate now for Bill Herzog's group to report. Their discussion was centered, at least originally, upon costs of evaluation.

MR. HERZOG: Thank you. Everyone was in a good mood yesterday. We had a lot of discussion, no real debates. Maybe we will get some today. I think I will introduce John Barson from the Office of Medical Education, School of Medicine, Michigan State University, who will make our report.

DR. BARSON: What the group attempted to do in examining the cost of evaluations is essentially arrive at some rational resolution of the issues that are involved rather than trying to establish technique, or even give priority to methods of costing. Three major issues came up with regard to cost factors in evaluating continuing medical education.

The first question ought to be the first one applied to a program, "Can we afford not to evaluate continuing medical education?" There, curiously enough, are two answers that seemed to come up in our group to that point -- one is "Yes" and the other is "No," depending on the viewpoint and who gains and who loses.

There are some areas of learning that apparently are not significant enough to warrant the expenditure of funds, or upon which the specific measurement of gain is not critical.

A friend of mine used to ask, "How thin do you want to slice the boloney anyway?" This, in effect, is the gist of the decision. If there are high risks involved in not evaluating, the cost can appreciably climb, and justifiably so. On the other hand, where there is low risk involved in not evaluating, it may be well not to burden the program with additional evaluation costs over minimal measures.

Another area of cost, identification of costs, and counting all costs has to do with the cost of ineffective evaluation measures. Inappropriate or inefficient evaluation measures obviously are more expensive than the effective ones, and yet quite often these costs either are ignored or somewhat concealed in the cost of teaching. Admittedly, ineffective or inefficient evaluation measures can double the costs which were cited yesterday by Bill Herzog, since we not only lose what it costs us to do the evaluation but also the losses of not evaluating appropriately are also added.

A third area of counting all the costs has to do with differentiating between the development costs of evaluation and the application costs of evaluation. If you ask any manufacturer or producer of material goods you will find there is an enormous amount of investment in research and development, the R and D factor. Research is often identified easily by laboratories and other sites and types of workers. The development is not quite so easily identified. The truth is that most educational institutions spend the bulk of their budget in development and a minor portion of their budget in delivery of what they have developed, but for some reason they are unable to articulate the fact they are developing instruction, they are developing evaluation, and these costs have never been identified. The models are seriously lacking.

This is my personal experience, the ratios I have seen with regard to development costs versus application costs in teaching-learning situations ranges anywhere from 4:1 to 11:1, which are very high ratios. This underscores the expense of ineffective or inefficient evaluation measures.

A fourth point in counting all the costs has to do with considering the amortization of costs from repeated use of proven evaluation measures. That is to say, the more often you use something, the less cost it is to you, especially in the case of evaluation where there should not be too much wear and tear, short of having to alter the program to meet minor changes in the personnel who are being evaluated, or for certain factors. This amortization factor is very seldom cited in accounting procedures, but proven methods can have a lower and lower cost as they are used more and more widely.

The fifth point under counting all the costs has to do with relating the evaluation costs to the available resources for the program and actual total program cost. What seems like an exorbitant amount for evaluation might not seem so much if we were to count the actual expense of a program which is being sponsored, counting all the costs related to it. This issue of relationship was brought up by a number of speakers yesterday. It was not explored much beyond that because we would have to get to specifics at that part, but this is certainly a factor that we would have to take into account in counting all the costs.

The summary of counting all the costs and the rationalizing of whether it is appropriate or not to expend funds in behalf of evaluation has to do with the question whether we are trying to do better what should not be done at all. This issue came up repeatedly in the discussion, that we may really be trying to improve on something which should be reduced. Again, this is another factor in cost accounting for which we must account.

The second major issue of evaluation costs, curiously enough, does not look like a cost issue at all, as has been cited earlier in the presentation today, and that is the use of medical practice as an efficient base on which to establish education evaluation; for instance, the medical audit.

Several points were made with regard to, "How do we select an appropriate evaluation measure?" Under that the first point might be that evaluation is appropriate if it is a part of his continuing education it probably should be closely allied with his practice of medicine. This suggests that the use of medical practice as an efficient base is justified. At least this was the consensus of the group.

The use of medical practice as a base for evaluation has several legal and ethical considerations. We are striving for efficiency, but the question is whether the cost of efficiency, so to speak, can be paid by people in the profession, or whether they are willing to pay it.

A third major issue with regard to costing evaluation has to do with the selection of the evaluation method itself, and three points were made under that by various members of the group.

First, and this sounds terribly obvious but an important issue to keep in mind, that the appropriate method is probably the least expensive over the long term. The sheer logic of that stuns you -- the appropriate method of evaluation is the least expensive over the long pull. Inappropriate methods, regardless of how inexpensive they are, certainly are not worthwhile. Just the garden-variety shopper knows this sort of thing, but quite often it is ignored in establishing evaluation programs.

A second point under selecting evaluation methods has to do with knowledge of cost estimates for evaluation, and these cost estimates are feasible and desirable, according to members of the group. Several illustrations of cost estimates were cited, one of them being one-third development costs, one-third application costs or instructional costs, one-third of the total budget applied to evaluation. This sort of ratio appeared in several projects mentioned yesterday. There may be other ratios. These are also variable costs, depending on how much advanced information we have about learners and how finely we wish to evaluate the experience that they have had in learning.

A final point in selecting evaluation methods has to do with a discussion held in the three sessions preceding this one, and that is the coordination of evaluation efforts with data collection regarding medical practice, the proposal being that evaluation would cost less if it could be a part of the on-going process of selecting data on medical practice so that we would not have to construct criteria and other measures to apply to certain learning situations. The results of medical care would be available to support or to change programs that presently are in effect.

A second part of that point has to do with the information exchange among evaluators, that it might be desirable to establish effectiveness guidelines among alternative means of evaluation so that each person approaching the evaluation decision is not working in an original area.

I would like to cite the one model I was responsible for introducing yesterday in summary: it is that in a sense the evaluation process that educators engage in, the decision-making, is not too unlike the diagnosis of a physician in working with a patient.

The physician has a certain amount of information about the patient when he enters his office. The more information, of course, the more accurate the diagnosis could be. In seeking further information regarding the patient, his condition -- in attempting to decide what sort of treatment to apply to the patient -- he has a certain number of constraints: economic constraints, logistics, whatever other problems he may encounter. He might order $1,000 worth of tests on a person to gain every bit of information he can, or he might choose to employ a much simpler approach, depending on the resources, and estimate of the seriousness of the problem. It is a decision-making process that the educator also faces. If the physician is able to deliver adequate medical care, with the large number of patients he is dealing with, there is no reason why educators could not make similar decisions without, as I said earlier, slicing the boloney thinner than necessary.

It is very difficult at this juncture of the discussion to say that I have covered all the points mentioned. I think Bill wishes to open the discussion to the rest of the group members who will kind of gloss the top of some icebergs we touched on in our discussion.

MR. HERZOG: Thank you, John. Are there other points that others in the group might have about the discussion?

DR. PERLMAN: I thought you covered our many, many points very well, John. There were two that you sort of passed through, and I would like to re-emphasize them because they do have an effect on costs.

The first one is that we talk quite a bit about defining evaluation and defining what we should be costing. One of the things that several of us brought out was that you have to decide what you are trying to prove or what you are trying to do with evaluations. If you go at it as industry may, or some type of very scientific research that wants to prove a definite point absolutely, we may find that it is very expensive to evaluate and we may never get there. Instead, if we can be satisfied to show some correlations and relationships between a particular educational program through an evaluation of it and seemingly some change in behavior, this may have to be satisfactory for evaluation, because of the many variables that we cannot control.

Tying onto that, Bill, you are really skipping something that you brought out, remembering Bill's paper from the other morning, all the costs and the very different percentages of costs for evaluation. Talking about what Hugh Petersen said about using evaluation of behavior as our major step, we felt we should not throw out all the other types of evaluation, even if we agreed that evaluation of a physician's practice will give us the best answer. What we should do is have several examples or several studies that try to match and try to evaluate the methods of evaluation so that everyone is not doing a survey which costs -- what was it? -- 250 percent of the cost of the project,

which I think is happening now. There are many surveys, and a great deal of money is going into them. Perhaps several groups with a very specific feasibility study idea ought to be doing surveys, versus clinic areas, versus practice to see if there are correlations. This is a little like what one of the earlier groups said about evaluating simulation versus practice. I think these two things are extensions of some of the things you had to say.

MR. HERZOG: Are there any other points that were brought up that others in our group would like to mention?

CHAIRMAN RISING: I was going to say, if nobody else wants to enter any further comment, that I would underscore something that John said. It really makes a difference what you are evaluating for, what objective you have in doing the evaluation. If you are merely evaluating in order to substantiate a request for more money from some granting agency it is too expensive even if it does not cost much. If you are evaluating as part of the educational effort in order to plan your program, in order to see how well you have done, in order to change course higher costs are justified.

As Dr. Clem Brown pointed out, if it is for program development, or actually if it is a major part of the education -- and in his type of situation the evaluation itself is the education -- the cost of evaluation is easily borne.

The formula that was pointed out, one-third, one-third, and one-third, we got to calling the "pie" formula in our group because the "p" was for planning, the "i" was for implementation, and the "e" was for evaluation. Literally both the "p" and the "e", both the planning and the evaluation are part of evaluation. You evaluate in order to plan, you then are justified in spending a fair amount of money. In reviewing some of our cost figures that I have recently gone over with Bob Neth, I think it is a ball-park figure that we probably spend two-thirds of our treasure in planning, using evaluation of needs and past programs, and evaluating current programs.

MR. HERZOG: I think there was one other thing we brought out -- and I do not know if we should bring it up because we did not get near a solution -- but in funding and costs John mentioned that you have to delineate what the costs are, and we spent quite a bit of time discussing long-range costs versus short-range, though we did not use that term. We may find that one of the methods that costs 250 percent to evaluate will make the whole education program cost much less over the next ten years, but it is a little hard to convince the funding agencies of this factor, and I think this afternoon we would probably all be pleased to hear solutions. We got as far as saying that that was a terrible problem, trying to convince them that some new method or some more expensive method of evaluation is going to save money over the next ten or fifteen years. They are looking at the tax dollars that come in this year -- period.

DR. BARSON: I underscore Jack Perlman's point for the fact that your costs can be increased by poor operation of other portions of the program as well, that quite often we are paying a high price for evaluation when actual improvements in the instructional program or the planning phase might reduce those costs, so these figures are hard to nail down to a firm amount.

I am almost under the impression that evaluation really should be a very minimal expenditure, if the preparations for it have been made appropriately in other parts of the program, that we may be paying a price for inefficiency elsewhere, like the evaluation study.

MR. HERZOG: Some points were brought out in the group that we really did not reach a consensus on. One was that we probably have not spent enough on evaluation in an intensive way, so we really do not know a lot about our measurements. We have a lot of guesses about the accuracy of the "reactionnaire" and what it might indicate toward behavior change. Several of us talked about the desirability of really intensively evaluating, trying to test all the tools and find out about how valuable our tools are before we are willing to cast out this tool or that because we do not know enough about our tools. So, maybe we are not spending enough, as in so many other areas.

Secondly, I think that we ought to have more centers that are particularly looking at evaluation, and maybe there ought to be more agreement among those centers as to how they were going to specialize so there will be a little more efficiency in their own approach toward evaluation.

CHAIRMAN RISING: Thank you. The summarizations have been excellent this morning. I have been very impressed by them. Knowing what went on in Bill Herzog's group, I know that John Barson did a marvelous job in condensing it, and I want to thank him.

The final group report, Group Five, will be by Hugh Petersen, the purist.

DR. PETERSEN: A little purity never hurt anyone.

CHAIRMAN RISING: Just so you don't overdo it! I believe in moderation in everything, even purity.

DR PETERSEN: Dr. Donald Korst will give our report.

DR. KORST: It is interesting to be last on this program and to have listened to the various summaries because certain things begin to come through, and I think as we sift this conference we are going to find that we do actually have some consensus of opinion.

Hugh Petersen made a good position yesterday in his written report on the needs for behavioral methods of measurement. We argued the position that the meaningful measures of performance change. Would the behavior of physicians in applying knowledge gained in continuing education ultimately influence patient care? The problems lie in how to develop the measurements of learning behavior. We discussed the logical reliability against the statistical reliability. The question arose as we approached the problem, "what are valid criteria?" Are these mere paper-and-pencil tests, or can we really go further than this? We debated the behavioral versus the nonbehavioral goals as the best methods for evaluation. We defined six problems.

The first problem was what are the methods that measure change in performance, because a change in performance should be an improvement. This is what we measure. The discussion included pre- and post-testing, vocabulary tests, assessment of individual performance, evaluation of hospital medical records, problem solving or case simulation, results of patient care, improved physician communication, utilization of selection patterns of laboratory tests, and hospital utilization patterns.

In application of any of the measures emphasis was placed on the importance of the manner that the methods are applied so that safeguards exist to protect the confidence and the trust of the individual physician. This has come out a number of times in the discussion.

The third problem: the communications between institutions, such things as the PAS program of hospital record auditing, and the medical audit program. Also, the problem-oriented case record, Larry Weed's system, was discussed.

The next problem, the fourth, was the attitudes of physicians to behavioral assessments which has been discussed at length this morning. Who should decide the behavioral criteria? We felt this should be a consensus of the physicians involved. It is very important to bring the physician himself into his own assessment and into establishing the auditing criteria or the medical center criteria. The success or failure of auditing would probably depend on the confidence of the physician. A functional set of criteria would therefore be developed if the men in practice were intimately involved in this process.

The fifth problem discussed was the popularity evaluation. Can we really go beyond this? Can we really evaluate short courses?

The sixth and last problem that we discussed was the motivation to participate in programs in order to be evaluated. In other words, how does one get a representative sample or an unbiased sample, because we are all aware that there are certain physicians who regularly attend continuing education, we have strong support, forty or fifty percent of physicians in some areas, but there is a group that we do not reach, so they are out of our sample of evaluation at this time. What are the attractions to programs? How can we stimulate this physician who is out of reach? Would the hospital medical audit stimulate the education?

The Minority Opinions. I think other members of the group may want to debate this, because I am not sure these were clear minorities. The first point: testing of vocabulary, pre- and post-testing, the paper-and-pencil tests in general were sort of a minority opinion that this was going to be adequate behavioral evaluation in the future.

I would like to ask other members of the group to make some comments.

DR. PETERSEN: I think, out of context, that is going to be a meaningless statement regarding the vocabulary testing because it does not seem to relate to anything. This came out of a discussion having to do with validating, using behavioral measures and relating them to nonbehavioral measures, and the comment was made that in the long run such a thing as a vocabulary test might be a simple, single, positive, predictive indicator of later success in some kinds of work that the physician would encounter. So, it was an example only.

DR. KORST: I think this is what we want because I am not sure it is fair to just arbitrarily divide these into minority and majority opinion.

DR. PETERSEN: I am in the minority all the way through this.

CHAIRMAN RISING: Don't be paranoid, Hugh.

DR. PETERSEN: What have I got to fear?

DR. KORST: The second point: the partitioning of continuing education activities to suit the individual needs of the physician; in other words, the very specialized conference. I think there was not a majority opinion that this was a good direction to go.

The third point: is patient care the only goal? Everyone seemed to think it was the major goal. At least we did not have strong feeling that there should be other goals than patient care.

The fourth point: the popularity -- or I think it was well put, the happiness poll of programs -- not much enthusiasm there.

The fifth point: this one really should be 1.a., or something in the majority, because I do not think we really decided minority or majority on this, but I think everyone felt there was a real need to know more about what individual departments or offices of continuing education are doing. There ought to be some coordination or some way of disseminating the information about what programs are being evaluated, how they are doing, to avoid unnecessary repetition in costs. That really should not be under minority. It is a separate thing.

The Majority Opinions. I really hope our group will discuss these. The first is motivational factors, which seem terribly important. These should be studied further. The achievement of improved patient care seemed to come out as a universally accepted motivation of physicians. The physician responds to this in a very positive fashion. Any evaluation programs that were related to the achievement of improving patient care were well accepted.

Second: sample size -- and this really follows with motivational factors -- in some way we ought to improve our sample size, the distribution of our samples. This is reaching to the physician we are not reaching now, and this is absolutely necessary to get a representative group with which to work. If we are going to get any kind of a random sampling of impressions and programs, it is necessary to broaden the sample.

Third -- we spent a lot of time talking about this: smaller programs related to individual physicians or to individual hospital programs; in other words, going out to the community. We felt this was a great need, to work with individual practices, individual hospital programs that involve the hospital staff, the hospital administration in a very direct manner, the hospital organization involvement toward (1) self-evaluation and (2) auditing. These two things, the education and the auditing, ought to be brought together in the smaller program where the physicians themselves participate. They may develop their own criteria, with help, with aids, with suggestions. The universities, the continuing education departments can supply the expertise, the advice, the education, but if the physicians themselves help in developing the program it should help the motivation. This involves the case record review, or improvement of case records, this also involves the hospital staff time investment and the developing of hospital education programs.

Fourth: some minimum requirements for continuing education. This relates back to number three, that the hospital auditing and education committee within its own organization can begin to develop a minimum requirement, not a punitive sort of approach, but a constructive approach that might point out individuals or direct individuals into certain kinds of continuing education where the hospital felt they had a weakness, through auditing or through self-assessment. The standards of records would be required in this program.

Fifth: cooperative or coordinated studies should be started. These would be selected programs through continuing education departments that would be developed to obtain a sample size that would allow random sampling to evaluate effective measures of behavioral evaluation as to reliability, success, cost, numbers of measurements, etc. The validation of the behavior needs a random study to include the comparison between the statistical probability analysis, the paper-pencil tests, with the behavioral evaluation that will affect, ultimately, patient care and hospital utilization.

We felt the test program or the study should be small enough, it should be simple enough for a single topic-oriented program that could be used in a number of centers for the evaluation. The physician participation was essential to get the physicians' self-acceptance of a program.

One example idea that was discussed would be a coronary care program where this would be tested as a uniform approach to a program, a uniform evaluation.

Finally, this is my own comment in summary to what we have heard this morning, it seems to me that the time is now. My own assessment of physicians in practice in community hospitals is: they are moving very rapidly in the direction of self-assessment and of auditing their performances. There is a very strong feeling among physicians that if they do not do it, somebody else will. I think this is an accepted fact, and that you will find a very receptive audience in most hospital staff organizations.

DR. PETERSEN: I would like to make a few additional comments, just to reiterate a couple of things. We talked at great length about behavioral measures and there was quite a bit of argument about whether or not that was really an appropriate approach. Someone would say, "well, does it not depend upon your objectives?" Since I have been around Steve Abrahamson for so long I answered that immediately. My question to you is: do we really have only one objective? We only have one objective as far as I am concerned, and that is improving patient care. We do not have an objective that is called "amusing physician students."

QUESTION: Don't we? Do we not also need to please and stimulate our students?

DR. PETERSEN: Again, I am naive, I guess, but I have been here as a purist, as I say, if we are only trying to do that, if that is our goal, then we have to have behavioral measures. Once we have decided to do that, we can take a slightly gentler course. We can say this: if we can come up with some definitive studies where nonbehavioral measures are validated with behavioral measures, where we can show very high positive correlation, then we have some empirical basis for saying that we do not have to evaluate using behavioral measures every time. The areas, that is, the domains wherein we do this, have to be very well described, very well planned, and definitive studies must be made.

Dr. Korst was talking about sampling, and it is a very tough topic. Perhaps I can now show you it was related to what we were talking about. Supposing we take all of the CME programs in coronary care throughout the country. If stratification within that domain is necessary we could do that, but let us just say we take that domain and we randomly sample ten of these and re-throw our evaluation efforts into those ten where the results can be, because of their statistical characteristics, generalizable, then the money would be well spent. Since we cannot use behavioral measures for every study, obviously, this other approach is to me the only logical approach at this particular time. So, it is a validation study, and the results could be generalized to a much greater population than we can at present. This requires great cooperation, and I am not sure where that is going to come.

As far as the rest of the group is concerned, do we have any comments here?

DR. WELSH: I want to congratulate Don Korst for bringing such order out of chaos. As one of the majority stated, all of the groups have articulated the fact that the prime motivational factor for most physicians in practice is improvement of patient care. To echo what Dr. Korst said about hospitals being ready and moving to develop a system, they see the handwriting on the wall. We all see that if they do not do it, someone else will do it for us, then actually the setting is ripe for the development of a good behavioral study in which they are going to develop their own criteria, whether a la Clem Brown or what other system they may use, they are going to develop their own criteria which will threaten them the least, which will be in-house, which will give the tools to them for improvement of their own patient care activities. They will need some help in this, and I feel the role of the educational institution is the resource to aid them in accomplishing this, and not to do it for them.

CHAIRMAN RISING: Thank you very much. We will have discussions of all the group reports this afternoon. I encourage you during the lunch time to consider what you want to put in the record this afternoon, the further points that you want to make, the debates you want to have with the purists, or the impurists.

THIRD PLENARY SESSION

CHAIRMAN RISING: We are going to open the floor to anyone who wants to comment about any of the group reports from this morning. I know that people were straining at the bit this morning, wanting to discuss things that were not part of their group report, and this afternoon I hope they are not too sleepy to remember what they were thinking of this morning.

DR. LESTER: Some years ago I spent an evening with the minister of health for the Kingdom of Nepal, and during the course of the evening he described his several years of study of public health in this country. I asked him, of all the things he had learned, how he intended to apply all of them. The sum total of his remarks was: "In my lifetime if I can achieve a single clean well in each village, I will have done more than I think time and capabilities will allow me," to say nothing of all the rest of the things he had learned about public health.

In some ways I think that I, particularly, in putting on programs for general practitioners -- or seeing that they are put on -- feel that, yes, we should measure in terms of behavior. I cannot see in the forseeable future how I can accomplish that for the State of Minnesota. Certainly some hospitals will have ways of measuring this.

One of the tasks I have, and I am sure certain others of you run into this, is an institution-preserving meeting, an annual meeting where head count is all important. Members must turn up. They must pay dues. They must meet one another for fraternal reasons. They must have a house of delegates meeting where they pass resolutions. The institution achieves many other things for the doctors and we think, rightly or wrongly, eventually for their patients. In this setting we are asked to produce a meeting, and I do not have imagination enough to foresee how I could evaluate changes in behavior of physicians from all over the State of Minnesota.

Like the minister of health from Nepal, is there some lesser method that I can use at this time to salvage some educational value from a meeting which I am not asked to evaluate. What can I do to improve the educational value of their meeting? What technique could I apply? I would like to ask those here: is the pre- and post-test still valid with those kinds of limitations?

MR. JENKINS: I have a question. Are you really after an educational objective, or are you just out to get people together for a fraternal gathering? If this is the objective, a head count is probably a valid evaluation, if that is all you are out to do. If, indeed, you are out to change behavior, then the only way that you can know that you have changed behavior is to measure the behavior both before and after to find out if it has changed. You have to go back to the objective. What is the reason for doing it? What do you think you want to accomplish? Once you have this down, it becomes a very clear task as to what you need to do. If your objectives are stated in what you want to accomplish, then you just go about the task of measuring them. It may be that the instruments do not exist and you have to develop them, but I think they can, in all cases, be developed with some degree of expertise from people such as those present here today. If you want to have a fraternal get-together, that is fine. There is nothing wrong with that.

DR. EISELE: I object to that "either/or," either it is a social gathering or an education experience which absolutely must be evaluated. You have another alternative there, an educational experience which will not be evaluated.

MR. JENKINS: May I respond? Can you prove it was an education experience if you do not evaluate it?

DR. EISELE: No. I do not believe I am always obliged to, either!

MR. JENKINS: Then you cannot call it an educational experience. You are only guessing that it is an educational experience.

DR. EISELE: I am not asking for a grant support, so I am not obliged to prove it except to the satisfaction of my enrollees.

MR. JENKINS: I do not care whether you are or not. Is this an educational experience here? I think it is. I have learned something. You are not going to measure it to find out. I think I could have by measuring some objectives, and we only had one, as I understand it, to write the report that is coming out of this. As a school teacher, sir, who was charged last year with evaluating youngsters, putting grades down, and you, I think, are involved in a medical college, are you not? You are going to give grades to students on what basis?

DR. EISELE: No. We do not do that any more.

MR. JENKINS: You do not rank any students?

DR. EISELE: No.

MR. JENKINS: No ranking?

DR. EISELE: No.

CHAIRMAN RISING: May I ask whether you are talking about evaluating the students or evaluating yourself now? Did you evaluate yourself as a school teacher last year?

MR. JENKINS: I tried to. I was very inadequate at doing the job.

CHAIRMAN RISING: Well, could you call it an educational experience? Is a high school education an educational experience or have we been kidding ourselves?

MR. JENKINS: For the teacher or for the.....

CHAIRMAN RISING: For the children.

MR. JENKINS: For the children? I don't know.

CHAIRMAN RISING: I think this is Dr. Eisele's point. He may be taking exception to people saying: "continuing education in medicine is not educational unless you do evaluate it," whereas for eons past they have not objectively evaluated any other educational experiences -- Socrates to graduate education in zoology.

MR. JENKINS: If we are going to make statements about what is to be gained, in other words, why did we do it, if we are going to ask a doctor to come to it, we should have some reason for him coming. If the only reason is that he will be exposed to someone's ideas, then fine, a head count is valid if this person did, indeed, present ideas. Someone is paying for that. That is the question that was asked this morning, and supposedly the physicians are asking, "who is paying for it?" Who pays our salaries? In essence most of them come back to the taxpayer at some point.

CHAIRMAN RISING: Some of them, incidentally, come from the taxpayer. That is mainly the RMP programs.

MR. JENKINS: No, all the people who work for universities.....

CHAIRMAN RISING: I take exception to that. The University does not pay my salary. The doctors who enroll in our courses pay many of our salaries, Dr. Eisele's included, I feel sure. 65

MR. JENKINS: All the doctors pay those salaries?

CHAIRMAN RISING: That is right.

MR. JENKINS: How much of the doctor's salary comes as a result of taxpayers' money to Medicare?

DR. PERLMAN: I would take exception to that. How does the doctor pay your salary and not the taxpayer?

CHAIRMAN RISING: Dr. Eisele's program, the Minnesota program, my program, the program at Albany, a good many of the large, maybe not educationally successful, but quite popular programs, are paid for by fee income. Doctors pay fees.

MR. JENKINS: O.K. I disagree, because most of the fees are not paid by individual doctors, they are paid by... it is not tax-supported, it is taxpayer-supported, citizen-supported, because it comes from the hospital till.

CHAIRMAN RISING: Now wait a minute. Money for our programs does not come from the hospital till.

DR. CLEMENT BROWN: To follow up on Keith's point. Where is the money coming from? Suppose it does come from the physician; suppose it comes from time out of his practice, away from his family. I think one could make a good case that we are really perpetrating some gigantic fraud on physicians, getting them to spend time away from their families, away from their practice and from the patient care they could be delivering to attend courses that we have no idea whatsoever whether this has any influence on their behavior, on their patient care, on anything, if we do not measure it.

I think we must be concerned. We may be perpetrating a gigantic fraud on the American public, the American doctor, and everyone concerned. I am not saying that we are; I am saying that this may well be the case. As a matter of fact, almost every time an educational experience is evaluated a reasonable kind of way, is not this exactly what we find? The New York television programs, the Butterworth thing, the McGuire-Babbitt thing that they evaluated on cardiac auscultation, every time we measure something in a reasonable kind of way regarding continuing education we find it has no effect at all.

CHAIRMAN RISING: Those things, incidentally, were not financed by the enrollees.

I would like to say one thing now that was said in our group yesterday. I have not heard it repeated here, I believe. That is, that the "evaluators" here probably go to their doctors and trust them to make what may be a life or death decision about him, but he does not trust the doctor to have a fair idea of whether a program is educational to him. I submit that, if a doctor spends his time and money going to an "educational" program, there is reasonable expectation that it is actually educational, and in that sense a head count has at least slight validity. Nobody would say that if you can economically evaluate the change in behavior, you should not. Probably anyone would agree to that. The "either/or" position and saying that we are "perpetrating gigantic frauds" are extreme and unfair.

The programs put on in your hospital, Clem, are an entirely different thing. There is no question that you are 100 percent right in a community hospital, where all of us will probably agree is where most continuing education should be. I have noticed here in this group that, in spite of having pretty good luck in getting a mix, we have had a woeful shortage of community hospital medical educators, and nonphysician health educators. We have had very little contribution about them, and all the time we have had to try to transpose "physician" or "doctor" into "health worker" or "health professional."

MR. HERZOG: I think the phrase "perpetration of a fraud" is a little strong. It makes a point, but I think the point is that you are saying that you have got to somehow justify the expenditure. If you can not justify the expenditures on continuing education how do you know where you are going? It is my viewpoint, unless somebody can change it, that I am not the least bit more confident that the money spent on evaluation has been well spent because I have never seen the justification. I have not seen any more good evaluations of evaluations (ones that really measure what they set out to measure) than I have for continuing education. I will take the devil's advocate role and say you have both got to prove this to me because I do not see the value.

DR. THOMAS BROWN: I wonder whether anything that happens to anyone lacks educational content. I think that all of our experiences are a cybernated kind of thing that, you know, we feed out and we get back. The question that we are addressing ourselves to at this conference is: "what is the role of evaluation in the educational process?" What can we make observable, and what ought to be made observable so that we can use the data from evaluation to improve our programs, that evaluation has a purpose. I think any experience is educational. It may be negatively educational, it may be positively educational, but it is educational. I think we have to take a look at, "what is the role of evaluation, the data that comes out of evaluation that we can feed back into our programs that can make us more effective." I do not think anyone really wants to disagree with that.

CHAIRMAN RISING: Incidentally, it may be related to how you can better intrigue professionals, health professionals, with your program, how you can get them involved. You can have a marvelous teaching instrument, but if no one uses it it is valueless. In the community hospital where there is a certain amount of clout, you can make people use it and they will begin to appreciate it.

DR. VANSELOW: I want to support what Bill Herzog said. Clem, you know it is nice to get up and make statements that none of these things have any educational value...

DR. CLEMENT BROWN: I did not say that. I said, maybe we are perpetrating a gigantic fraud, I do not know.

DR. VANSELOW: The thing that concerns me is our measurement devices. Are our measurement instruments accurate enough really to measure what we think we are measuring? That is one thing that really bothers me.

The other thing I would like to do is get back to John Lester's comment. John, I really do not think we can expect you in Minnesota with the Academy of General Practice to do the type of evaluation we have been talking about here, nor do I think we can expect the great majority of people who are doing continuing medical education to do it for every program. I would like to see some studies done from which we could generalize so that we can say that, if a program is put on under the following circumstances, it does or it does not have any impact. It seems to me that if you could do that, if you could put the money into a few studies and then generalize from them, you would be in better shape. I cannot imagine, however, how we are ever going to be able to do the meaningful type of evaluation on all of our activities.

I wonder, I think Hugh Petersen suggested this this morning, whether this is not something that this group ought to recommend: that a cooperative study be set up where these various techniques can be evaluated with the hope that we can generalize and not have to evaluate everything from there on.

DR. LEMON: It has been curious to me today and yesterday, as we have been talking about evaluation, that we have frequently been moaning that there are no good studies to hark back to, but no one up to this point has mentioned the fact that at least a fairly well sketched out study was done right here in your bailiwick entitled "The Epidemiology of Continuing Education" by Chuck Lewis. When I first read it -- that was one of the days I spoke about this morning when I thought I had better look for a place

to practice. After twenty-five years, would you like to make a few well chosen comments, in four-letter words, about this thing which demonstrated that you had not made any impact in Kansas for twenty-five years?

CHAIRMAN RISING: That study was not done behind my back or without my knowledge, and it was not published without my O.K. I read it in manuscript and I argued with a few of the statistics, because he used the data differently, I thought, than they probably should have been used. This was not a question of measuring behavior and end results, I think Chuck would be the first to tell you that Kansas was one of the few places in the world that he could have got the data to do such a study, because we had been doing enough evaluation of at least a "market evaluation" type -- and this is the term he used -- to make this much possible. It would be difficult for me to prove to Dr. Petersen that we had changed behavior of doctors in Kansas. I think it would be difficult even if we had had some pretty sharp instruments because the field is broad, both geographically and in the number of persons involved, and there are too many contaminants.

I have said, and I would like to say again -- I have said it on the same platform with Clem Brown at Wes Eisele's place at Estes Park when he was having the program -- that continuing medical education is not a course or a set of courses. Continuing medical education is something that goes on continuously and in many different ways in each instance. It involves, as Don Williams, who headed continuing medical education in British Columbia many years so aptly said, work in the doctor's own study, in his meditation, his reading, it involves consultations with colleagues. It involves curbstone consultations as well as formal consultations. In many, many ways learning goes on. For continuing education to be truly effective it needs to go on at the home base of the doctor or health worker, which means generally in his own community and in his own hospital. It needs to be an everyday thing, not a course thing.

I would like to repeat an analogy with religion that is old. I see the role of *most "courses" -- and this includes the courses that the organizations and institutions* put on and the voluntary health agencies put on -- as analogous to revival meetings. People go to a revival meeting for various reasons. They go to feel uplifted. They go to courses because of intellectual uplift. This is not necessarily directly educational but it might indirectly be because, as in the case of religion, it may influence what goes on when they get back home, when they go back to their own community, their own hospital, their own home.

If we inspire people to study a little more, to have more consultations, there is no way that we can, within any reasonable cost, evaluate the change of behavior because we may be talking about the use of diuretics, but it may inspire them to get involved in something else about hypertension and into stroke and into rehabilitation by a circuitous route. How do you evaluate that?

This "either/or" business does bother me, and to think that all continuing education occurs in courses or in relation to RMP programs or to medical school programs is obviously fallacious. There was continuing education before there were schools. It has gone on since at least Hippocrates and probably before that. What they taught might not have been very good, but I am not sure that what we are teaching now is any good. Five years ago it was great, but we know today that most of it was pretty lousy. If we had tested and saw we had changed behavior, we might have seen that we were changing behavior in the wrong direction. This has been mentioned many times. What are your criteria? Five years ago cyclamates were o.k. and tolbutamide was great, but now they seem to be out!

DR. PERLMAN: I am interested in something that one of the groups brought out this morning, but we have not explored. There are tools available, I think you just said that, but we cannot neccessarily make sure that the tools are used. We have to figure out how to get the average community hospital and the average physician to be knowledgeable of these tools and to use them.

I would like to ask the group to explore what role the university, the medical schools and postgraduate education -- not just the postgraduate education groups that are represented here, but the medical schools themselves -- have in teaching, and how they can go about teaching physicians and medical schools to use the tools, such as medical audit and self-evaluation.

DR. KORST: This is a very important point. It seems to me that continuing education is really a frame of mind, and this is something that needs to be instilled in the student at the time he is a student to continue on in practice. One of the most important things I think we have talked about today is not only getting to the physician in the community hospital -- and I say this because I am a director of an education program in a community hospital but I am associated closely with the university -- but in the medical school we do not do this, we do not teach students about how to continue their education or the importance of it. We do not show them ways to evaluate, although students are beginning to get more into self-assessment programs. The students now tell us that they want the pass-fail system but they want examinations. They want to evaluate themselves once or twice a year, have us show them, with them, where their weaknesses and their strengths lie. They do not care about the grading of this, but they are very anxious to have a self-assessment.

About auditing, I think this is just an addition. It would be all right to talk to medical students about the medical audit in terms -- maybe Dr. Brown will take issue with this -- but I think a great deal more has to be done before we can say to medical students, "now this, in effect, is something you are going to be working with." This has to be proven. My own feeling is that it would be very acceptable and very good, but it has not struck me that this is generally accepted by the medical profession as an ample way of evaluation. I think it should be looked at but I do not accept it without reservations.

CHAIRMAN RISING: It is not really accepted by many community hospitals!

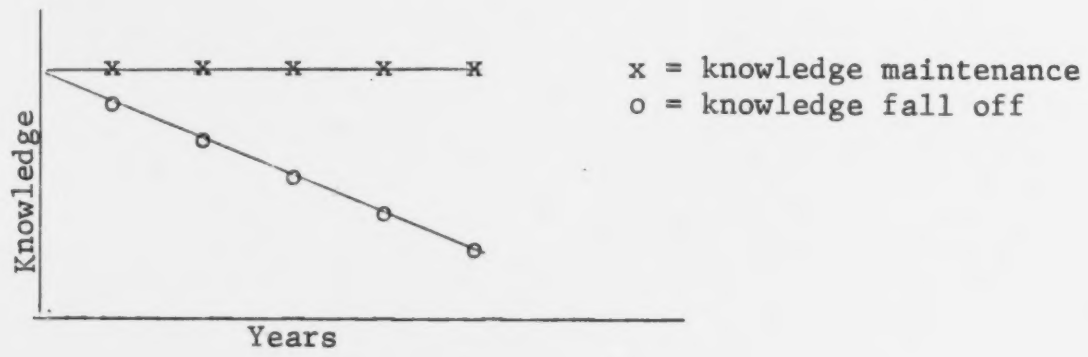
MRS. SHORT: I have two points to address myself to, one being that I am participating as a representative of Vanderbilt at this time, but prior to my association with Vanderbilt I was associated with a community-based educational program that became firmly established and grew through the evaluation of the doctors who attended, who repeated their attendance, and who verbally expressed themselves to the effect that they came back because what they found was quality and practical application in their practices at home.

Whether it is done formally or not, evaluation goes on all the time by the attendee and by the program producers, and because of the increasing social problem and the increasing social interaction that is making this more of a public problem, we are going to have to have more concrete tools, but you cannot get away from the fact that evaluation exists. We are not putting it to the best use, however, and to the most practical and useful methodology where we can more concisely employ it.

My other point that I was interested in expressing was that it seems to me that a lot of the comment has been directed to evaluation as something that occurs far along in the process, kind of after-the-fact rather than before-the-fact, and to me evaluation is part of the total process whereby you may establish a well-defined, concise goal, and as soon as you have that goal established the intermediate steps fall into place and still leave you leeway, but the evaluation is built into the program; it is not tacked on at the end.

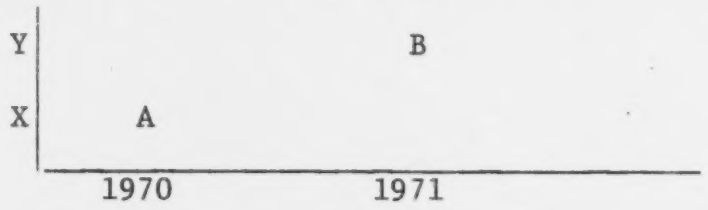
DR. FIFER: I would like to say if we use only behavior as the means of measuring success of our efforts, we could be misled a little. I talked with John Lester about this earlier.

This is my knowledge of, say, infectious mononucleosis that I know something about right now. If I do nothing about my own maintenance in infectious mono my fall-off curve is like that (see graph). If you measure the change in my behavior, at the last point my behavior will have changed negatively.



If I go to enough courses on infectious mono where I perform what might be called "knowledge maintenance," assuming that there is nothing new, then my curve is going to be like this, and I am going to be taking care of patients with infectious mono five years later at the same level as before. If you measure my behavior it will not have changed. If I did not go to those courses and you measured my behavior you would find it deteriorating. I do not know how you are going to take care of what occurs in knowledge maintenance, and there must be quite a bit of it that occurs.

Another thing I think about behavior is that we are about, because of the pressures of the moment, to give a course in the use of L-dopa in Parkinsonism, a disease which is a terribly important thing for practicing doctors. Because L-dopa has suddenly become available it is an important thing, it is a one-shot, and if we are at point A in 1970, and you do a behavioral analysis of what doctors do with Parkinson's disease today, you will find that X many use this new agent, and a year from now you will find this new agent being used by Y many doctors, whether we have taught them anything or not. They are going to use L-dopa, not because of a course I give. There are other places for them to get good advice, but they want our help and our reinforcement of what they get from other sources. The doctors are going to be using the drug whether we give the course or not. We would like very much to have them apply it well.



This is an example of two behavioral measurements, neither of which may have anything to do with the course I have interjected; in fact, one of which could be a reverse correlation.

DR. PERLMAN: Two comments, now that Bill Fifer has spoken. There is a definite measurement, particularly in your first example, Bill, because the important thing in evaluating, self-evaluation and medical auditing is to mesh against the patterns of care that you expect. If you expect that the ideal care for hepatitis has not changed, that is what you expect; therefore, you have measured the behavior. Merely making the assumption that knowledge will drop off if you do not do anything, if you do not use something, a relationship implication can be established that we have done something. It will not be negative.

I really wanted to talk because I am very concerned about the thing that Dr. Korst said. If I understood him correctly, I find myself falling into two possible conclusions, and they scare me. The first is that we ought to be very careful of these immature medical students because they cannot evaluate. Until we have proven that self-evaluation

and medical auditing is any good, we better not expose them. I am probably being a little strong, Don, but you could get that implication out of what you had to say, and that does bother me.

From a personal standpoint, because of what I do, I think it has been pretty well proven in many hospitals, and Clem Brown certainly is an example, that medical audit is an effective tool in hospital evaluation. The first thing that concerns me even more, whether it has been proven or not to anyone's satisfaction, is the fact that it does work. I think we have to expose our students to the possibilities of something that is not just blue sky; it is something that has worked in some places, that certainly can be improved, as anything can be improved upon, and teach them this concept.

I do not know if you were trying to stimulate or what because you are the one who brought up self-evaluation of medical school originally.

DR. KORST: I was hoping I would get some comments like that. That is why I said it. What I really said, or what I meant to say, is that I am not opposed to the students being taught this concept. I think the students ought to look at this as a research in health care, just like we are teaching students to work with computer interviews. I think the students would take hold of this and work with it very well. I do not think we should present this to the student as a matter of fact "this is how we are going to evaluate." We have to investigate and study it. I did not mean that the students should not have it. I think they should have it but in a way in which we are looking at a lot of these programs.

DR. PERLMAN: One quick answer to that. I do not see it as being an "either/or." I do not think it is the same as computer interview, because computer interview is not available at this time from a practical sense in any community hospital if it is not their own experiment. That is true with medical audit. Medical audit in some form or other is probably available in fifty percent of the hospitals in the United States. The tool is there. It may not be used well, but the tool or the mechanism can do something. It is there, and it is somewhere in the middle of the two.

DR. PETERSEN: Let us not misunderstand the use of the behavioral measure. In using it we do show a difference, and we do know what the difference means, that the physician is behaving differently. That was the point. Now, we can speculate about WHY, and of course we have a controlled study which he has used, non-experienced or non-continuing experience with continuing experience, operation defined as "continuing education courses," or attending these courses and not attending the courses. It just injects a note of certainty into what we are doing. That is all.

DR. LEMON: A couple of points about behavior which I think were stimulated by this, and I had hoped that this might come out of my question about Chuck Lewis' paper, because it would seem to me in reading it, that since we were measuring in fact an end result, I have forgotten the details, but it was deemed that the impact was "non-successful," because no change had been observed. This points out the fact that maintaining the status quo may in fact be an objective that is worth reaching. These individuals who did not improve in Kansas practice also, perhaps, did not deteriorate over this period of time as a result of their exposure. That might be worth something.

The other thing I think about behavior, or at least the effect on patient care, which I personally think is the only reason for being in this business, nevertheless the measurement of it does not have an invariable relationship to knowledge. I think we should keep that in the back of our minds also, because there are several reasons why people do not behave to a certain standard, and they are not all due to the fact that they do not have the knowledge. That has been said in several different ways, but it does not hurt to emphasize the fact again.

We have a delightful person in my part of the country who tells the story about talking to a farmer, who, in turn, had talked to one of those agents who drive around -- what do you call them? -- a farm extension agent. He had made the effort to get a

farmer to improve his farming practice in a particular area, and on several occasions he got into discussions on specifics with him. About the fourth trip around the fellow said: "Listen here, young man, I already know how to farm ten times better than I am doing." So, he knew better how to do the farming but he was not doing it.

CHAIRMAN RISING: I am sure you wanted a little more comment, and I would not argue about the conclusions that Chuck Lewis came to in that paper except to say that he was looking at a rather narrow area, and was not evaluating the type of thing that I would like to evaluate. He was evaluating from a point a view of a public health person. It is just conceivable, and we have mentioned several frauds that have been perpetrated, it is just conceivable that if you evaluate an educational program and find that learning has not occurred or that behavior has not been changed, it may indicate there was something wrong with the evaluation. Perhaps you were looking at the wrong things or using the wrong criteria. Evaluations probably are not always perfect.

DR. BARSON: I am getting apprehensive that somehow the proceedings of this conference report will be something like: "These are the techniques which should be used in evaluating medical education; therefore, if you are not using them you are really in the wrong ball park." I glean from a number of comments that there is more to evaluation than the instruments possibly have to offer, that eventually, like the term "the unexamined life is not worth living," basically evaluation ability is a personal attribute which you try to stimulate or enhance in a person, and it should not be confused with the measurement techniques or devices or the indices which we have established.

I say that in light of the fact that we are not even certain which evaluation instruments to use, how often, or when. Typically, we make it a terminal activity, when actually we all know it is a continuous activity. Perhaps we should not offer an examination at the end of a course any more than you offer the measurement of a person's life worth after its fifth year. I mean, it has not lived its entire term out. It may require a sequence of experiences and, therefore, the measurement at any one point may really be irrelevant to the person's potential.

I am worried about that because I think, if anything, I would like to see in the proceedings that develop here this essence that evaluation is a response to an experience, both to the individual and to the observer, as opposed to the pure instrumentalities that we have to bring.

I think the funding agencies that back up educational programs are tending to be misled along this line and, therefore, are satisfied with a minimal expression of evaluation. I would call for more data on evaluation -- not just more instruments listed in the appendices somewhere -- but largely what sort of behavior can you expect from the participants, both the staff and the learners. I am very worried about this as I see ourselves constantly battling over the instrumentalities.

DR. BANK: Several times medical schools have entered the discussion, and my own particular point of view is to support Jess and his slogan of "Forty Years, Not Four." My point here, and I think, Jess, you are the one who at least coined that phrase, if we are talking about continuing education we are talking, by definition, about something that is continuous. It does not stop at graduation from medical school or from any school. It is part and parcel of a profession.

I take, as a definition of a profession, one that sets standards that are higher than those who receive the services of the profession have a right to expect, which means, if you are a physician, it is your obligation to continue to provide the best possible service. In this day and age when Mr. Oppenheimer came out with a grand and glorious figure that the half-life of information is about eight years -- which we know in the sciences is even shorter than that -- there is a problem. Change comes about. It has to be a continuous process. Education has to be a continuous process, and it has to be a responsibility of educational institutions to provide for that continuing education.

There is one medical school, which will go nameless as far as I am concerned, in which I have talked with their people and they say: "We are not interested in continuing education one bit. We see our role as working in research and teaching medical students, and when they get out of here, that is it! We are not concerned with them further." I cannot subscribe to that point of view of an individual physician or any individual who says: "Well, I have now been graduated. That is it. I know everything there is to know." That attitude cannot exist in this modern society.

To get back to my original point, I do not know how this philosophy can be ingrained in medical school teaching. The point was made about PAS -- whether or not PAS should be put forth as a technique may be questioned -- but can you argue against having the principle of something, however imperfect that measuring stick might be, to use as a guide to go on and have some basis on which to try to improve yourself?

One other thing, while I am on somewhat of a soapbox: can we, through whatever evaluating techniques, say that we have in fact improved patient care?

During an earlier discussion a statistic was used: "Well, you know, we cannot say that we have increased the longevity of American citizens through all of the millions of dollars that have been spent in continuing education." This is the wrong statistic to use. If we wanted to use that as a measure, let's do away with automobiles. There's a way of increasing longevity. Someone used pneumonia as an example. You can have the best possible care for pneumonia, and if that is the thing you are looking at you should only look at that alone, not all of the other things that can influence patient care.

One study I am sure all of you are familiar with, one that was done by Katz relating to the introduction of a particular drug, it was a highly controlled thing, and, the point being, this new drug, an antibiotic, was not being used in the community. They had done everything, they thought, to get this drug introduced into the community. All of a sudden they found a sudden upsurge of prescriptions that were written for the drug. The reason, they found, they were suddenly getting this increase of prescriptions was the fact there was one physician in town who, unbeknownst to the researchers, was the one that everyone else looked up to, and the minute he wrote the first prescription for the drug, everyone else followed.

These are among the contaminants to which you referred, Jesse. There are millions of them -- don't hold me literally to that number -- but there are many, many of these kinds of contaminants that you never know.

Probably the greatest definitive study on the flow of information is in the story of hybrid corn. There is probably no open-pollinated corn grown to any extent in the United States. It is all hybrid corn. There are about seven steps in this whole process that have been identified.

There are those who are early adopters, there are those who are late adopters, and there are some who hardly ever do anything at all.

DR. ROYER: I think this is an appropriate time to present a thought that I have developed over the last eighteen hours. I would like to speak on the topic of physician attitudes toward evaluation of continuing education. I briefly raised this question yesterday in our group discussion, and sent it up as a trial balloon, but the importance of this has grown on me during the last eighteen hours and I would like to pursue it.

Several persons have spoken of physician resistance, and I gather there is consensus that such resistance does exist. However, I submit this resistance to evaluation; the resistance to continuing education factor is more complex, more deep-seated than we have described.

Specifically, I question the interpretation, and perhaps the validity as well, of Neal Vanselow's second assumption in his paper; namely, the assumption that the practitioner is genuinely interested in the quality of medical care he delivers. I suggest this falls in the "motherhood-flag" category, and any physician facing a list of top ten words, as was presented in this physician study in Michigan, would be amiss not to include those items related to quality medical care. "Yes," the Doctor says, "I am certainly concerned about how good a job I am doing. I certainly am concerned about keeping up to date." Accordingly, then, I question the interpretation of these data. The questionnaire does not indicate what priority these concerns enjoy in the physician's life and in his practice. Herein I find a real discrepancy. I submit that among his priorities, concerns other than upgrading quality care frequently overshadow desire to keep up-to-date medically. Some of these have already been alluded to: golf, stocks, politics, etc. I think the discussion topics that you witness in doctors' lounges in community hospitals would surely give another indication of peripheral concerns, concerns other than health care.

I suggest, then, there are more complex reasons, some of which have come up in our discussions (namely, reasons related to a political-social conservatism), which resists change, and unless we admit that such concerns -- these deeper ones, more complex ones -- often take priority over that of upgrading health care, we are at a loss to explain the gap between the questionnaire results and actual medical practice.

Now, a couple of examples, we find the physician paying lip service to on-going education who continues to use chloramphenicol or aminopyrine, and with the explanation: "I have used these drugs for the last five years and I have not gotten into any trouble. I haven't seen any bone marrow depression."

Another example, we find the physician subscribing "yes" to improved health care, yet very easily writing off all alternate systems of health care delivery -- Kaiser Permanente, Medicaid, and closed panel practice.

So, I underline the gap between the goals stated and the actual performance, and this is the same gap that we see between test performance and actual practice. A discrepancy between management of simulated patients, for instance, this came up in our group discussion, simulated patients, whether it is mechanical patients or whether it is a paper-pencil simulated problem, and the actual practice of that physician.

Example: The treatment of a sore throat. On paper, of course, we would all get a culture. In practice how many times do we actually call on the phone, "O.K. 250 mg. q.i.d.," without even looking at the sore throat?

Regarding physician attitude, I hear a note of pessimism, and unfortunately, or perhaps fortunately for patient care, the sense of urgency that we have talked about is echoed in the various media.

In packing, I threw in several things. I was reading all my spare time these few days, one of which I was delighted to find has three or four articles on health care. It is the Saturday Review for the week of August 24. I assume from the nodding that some of you have read these. Senator Ribicoff talks about the infant mortality and how our health care does not compare to other programs. Dr. Knowles then addresses himself to, well, it is really a broadside against organized medicine. Carl Cobb, writer for The Globe, talks about the shortage of doctors and what you find and makes a plea for compulsory insurance.

I would like to mention a few things that bear directly on this point that I present. First of all, Senator Ribicoff says: "All that keeps the medical care system afloat is the fact that millions literally have no knowledge of their medical needs" -- our present system.

In the Knowles article, what he finds objectionable in organized medicine is what I am trying to describe for the individual in practice, too many individuals in practice. Several quotes, please: "But we are conservative, we doctors, and we do not want change and we do not want to face certain facts."

The broadside at AMA: "The AMA has resisted every major social change in medicine over the past fifty years, since, incidentally, 1916 when the AMA was on record favoring compulsory health insurance."

If organized medicine continues to maintain its posture of negative vigilance he fears destructive confrontations of an angry public and ultimately the complete socialization of medicine in America. He says that, however, much as the physician wants to preserve his autonomy and self-determination, he will be forced to surrender some of his authority and work with others.

What I am saying here is that the "sell" we have been talking about, this desire to sell evaluation of continuing education to physicians, I find myself quite pessimistic about. Perhaps we are naive in that we think we can sell this type of approach. I come back to some of the comments that have been made about a broadside approach where we try to remodel the whole system, perhaps priority being at the medical school level where we work for this attribute of on-going education that has been mentioned by several of the speakers.

CHAIRMAN RISING: I would question whether we have been talking largely about how to sell it to physicians or how to sell it to educators. I would judge that the educational establishment is at least as cupable as the medical establishment which, incidentally, is not really represented here very heavily.

DR. FIFER: I do not want to refute a lot of points but, unless there is someone here from Chicago who can successfully contradict me, I believe that Christine McGuire has got data from practicing physicians which indicate that a PMT, a patient-management type of simulation on paper, does actually bear a very high correlation to what physicians do in practice. I believe this was done among practicing physicians, I think, in Rockford, Illinois. If anyone knows the exact reference I would be glad to hear it. If we did not have to keep measuring behavior every time it would be fine, if we could go back to a patient-management simulator on paper, it would be fine, but I believe that Christine McGuire has this information, not only about how medical students and house officers behave, but how practicing physicians behave. I would like to hear some evidence that this gap exists between what we say we do, a la throat culture, and we do. If there is evidence that there is a gap, I would like to hear it.

DR. PETERSEN: The generalizability of your results is directly dependent upon your sampling techniques, your sampling procedures, and if you do not have random sampling involved, then you cannot generalize your results and it would not matter what Christine McGuire says.

MR. COOLE: Dr. Fifer, it is unfortunate that these studies are based on a very limited sample and are not, in fact, generalizable. They indicate a trend toward a high degree of validity between actual practice and patient-management problems, but as of yet there is no generalizable study available in this area.

DR. PERLMAN: Our organization has done probably twenty or thirty studies which will generally point out this trend. There was one that I was involved in that I think was a good example.

We did a study on use of antibiotics in patients coming in for routine surgery; namely, hernia, appendectomy, hysterectomy, cholecystectomy, and hemorrhoidectomy. Of course we did not do a survey to ascertain what physicians felt they should be doing or thought was the right practice, but according to the learned scientific articles coming from all the medical schools you use anti-infectives infrequently.

Across the country, of half a million patients, so we do not have to worry quite as much about sampling validity -- I am trying to remember some of the figures, and this is teaching and non-teaching hospitals, very little difference -- nearly 20 percent of appendectomy patients; 48 percent of cholecystectomy patients, and all this type of thing. I think the gap that Dr. Royer is talking about does exist in many areas. If this information can be challenged it is on the basis of under-recording, so it might even be worse than the records show.

DR. IRWIN BROWN: For a day and one-half I, as a paranoid physician have heard my attitude and my moral fiber and everything attacked, but the subject is Evaluation of Continuing Medical Education. We can focus on the physician because he is vulnerable. I am serious about this. We can talk about him. We can say that more kids die in the first year in the United States than they do in Denmark, or wherever it is, but let us be sure we are applying the same measurements. Let us say we are counting all of them that were hatched out as being dead, or whatever, and count that into our statistics and not wait until ten days later and start counting the survivors and apply it to another kind of thing. We have got to talk about the same thing.

CHAIRMAN RISING: Excuse me, are you saying that they measure infant mortality differently in Scandinavia than they do here?

DR. IRWIN BROWN: Yes, they do, and then we come up and say we are terrible, our babies die and the others do not. There is a difference in how you count them. If a kid is dead the ninth day he does not go into some countries' vital statistics, and if he is born dead here he goes into our neonatal deaths.

We are sitting around trying to see how bad these doctors are, and none of them is motivated well. Of course a doctor is getting paid for taking care of sick people and his motivation in some cases is money, but I like to think in a lot of cases it is patient interest. He is concerned; at least that is how he got into it. The educator seems to think that he, the educator, is the only one who is well motivated. Every one of us is here because he is getting paid to do this, so it is money that is doing it to us, too. Let's get that straight!

Now, how are we going to worry about evaluation in continuing medical education, which is what we are trying to do, if we are just going to measure the doctor. I have not heard anyone say, "let us have the physicians as a group get a committee together to evaluate educators and see how they are doing, and study their motivation and their monetary interest, their interest in getting grant money. I think we are all people. Doctors are not gods that are fallen; they are just a bunch of poor people walking around trying to make a living like you are, really. I think we really need to get into not just talking about one segment of the health care field and his motivation. What is your motivation?

MR. JENKINS: I am an educator, not yet in medical education, although it looks like an interesting field.

I am in sympathy with what the doctor is saying. I said this yeaterday, and I think it should be in some of the notes that were presented this morning, at least in the notes that were taken yesterday, that we are hung up on one thing, and that is hanging the doctor up here and taking a shot at him, and you know, this makes a good whipping boy, particularly when there are not many of them here to argue back. Unfortunately, the only way we can measure the effectiveness of the educational program is to see, indeed, what happened to him, because he is the recipient of the activity.

If I may use an example, and so that we take it out of everyone's hang-up, I will use your children, since I was a high school teacher. What happens to your children when a teacher does certain things to them in the classroom? They go in, they teach, and the children are supposed to learn, and they evaluate, and they assign a mark of some sort -- pass, fail, A, B, C, D, F, whatever you want to use. To tell us how effective the teacher is, how effective the educational program is, we must look at what

has happened to the recipient or the student. Yes, I can say, "this teacher is good because he presented this fact and this fact and this fact, and he used these visual aids, and he did these things, he went on these field trips, and all of these activities." If that child's behavior is not changed, he did not do a damn thing! He wasted your money. He wasted my money.

Gentlemen, I have to come clear back to the original argument. I do not care where the money comes from to do the job we are doing. I do not care whether it comes from fees that doctors pay because they are convinced, you know, we are nice guys, so they are going to give us money, or whether it comes from taxpayers' money. It's money. It's resources. It's part of our gross national product that everyone is concerned about right now, and it is causing inflation and the whole rest of the thing. We have to be concerned about "do we get our money's worth?"

I am not worried -- well, yes, I am -- about being able to prove that we did what we want, because if it cannot be proven I say "wipe it out!!" If they cannot prove I am doing my job, wipe me out because it is your dollars that are paying my salary.

CHAIRMAN RISING: Excuse me, did I hear what you said, that <u>they</u> cannot prove you are doing your job, or that <u>you</u> cannot prove you are doing it?

MR. JENKINS: If it cannot be prove....

CHAIRMAN RISING: If <u>you</u> cannot prove it?

MR. JENKINS: If <u>I</u> cannot prove that I am doing the job I am supposed...

CHAIRMAN RISING: You changed "they" to "you," or rather "I."

MR. JENKINS: O.K. If <u>we</u> cannot, within the situation, prove that we are being effective, then wipe it out. It is not worth it. It is a farce and it is a fraud, to use Clem Brown's term, and that, as far as I am concerned, is a valid term and I will accept it. I am using it, too.

DR. CLEMENT BROWN: I said "maybe." I really do not know.

MR. JENKINS: If these things are true, every time it is measured we find out it is, and if these things are true then it is a fraud, because someone is being "conned." Let us look at it. I don't care, you know, we are all worried about, "can we convince others, and can we convince ourselves that we are valuable?" Yes, we are valuable as human beings. We just may not be doing the job right. What is wrong with that? We fail. We cannot all be winners, but we can all try to do the job right. What I hear in here is some people protecting, some people attacking, and nobody being willing to sit down and say, "O.K., now is there a problem?" First of all, let us determine, "is there one?" I think I hear there is one but maybe I am biased. I am an attacker, I know that.

All right, let us first of all determine, "is there a problem?" If so, how can we correct it and what, then, becomes our objective? I do not think we are even talking the same terms when we talk evaluation. The question I wrote down this morning is: "What is evaluation?" I'll bet we could not all write down a definition that was even similar. I question whether we could come out of this meeting today and put on that board a definition of evaluation that we all can agree to. If we cannot do that, what have we done for the last two days? I ask you, "what is evaluation?" I do not know. I am an evaluator, but I do not even know what it is.

CHAIRMAN RISING: You have proved that you have done it?

MR. JENKINS: No. I have not. You have accepted that I have proved I have done it.

CHAIRMAN RISING: I didn't accept that, and there are some others that have expressed doubts.

MRS. SHORT: I would like to make a brief comment on Dr. Irwin Brown's comment. Our main activity for the last several months has been to ask the physicians to participate in criticizing what we are doing, what we are putting into the program, how we are presenting the program, whether we are taking it where it is needed, and what would be of enough interest and inspiration to them, offered in the most helpful way. If we could get them to be as critical of us as we are of them, perhaps we could thrash it out and produce something more useful.

DR. LEMON: I just wanted to point out that our discussion group, as you know, took as its major point that we thought we ought to get the practicing physician involved in evaluation of continuing education as well as his participation. That is the same section that Keith Jenkins was in, so we are concerned about that. I believe it is very important; we thought it most important.

The second point I want to make is in my area -- at least my experience -- is much like Marilyn Short's. We are very anxious to get physicians to do just that. I think we are making progress but it is an uphill struggle, and I have some sympathy for it because I have been in practice. The reason is not that they are not interested in continuing education or our program. It is because they are overwhelmed with other things. It is hard to get practicing physicians as involved as we would like.

DR. IRWIN BROWN: Here, at least, Clem Brown and I agree. For years at the State Medical Society meetings I have maintained a booth to ask the practicing physicians for their opinions and inviting their comments and criticisms. I am sure we all do something like that. That is part of your evaluation, but what I was talking about in the last day and one-half -- and I was not in the other group meetings and did not get that out of the group reports -- but in the last day and one-half we have talked about motivation and other behavioral changes in physicians, and to my knowledge no one has been saying anything about motivation and behavior changes in the educator-evaluator.

As a physician, I wonder what I would do if someone came to evaluate me if I knew he was an evaluator, if I knew his motivation, and if I knew what I was going to do to change his behavior. If I thought he was going to evaluate me with a good motivation, whatever that is, and it was going to change his behavior so that he would put on a better program for me, I would donate my time. To my knowledge this has not been part of our subject material as a motivation behind what we are doing. You know, good patient care, everyone can get behind that screen and say, "good patient care." As a physician what concerns me is, "why is this evaluator motivated to spend his time checking into me?"

DR. PERLMAN: From what Irv Brown said before, I think I was slightly misunderstood because we are equating two things which are not necessarily equatable. Dr. Jerry Royer will have to defend himself, but I am defending myself because I think it is important for the understanding of the group here, not because it bothers me.

As a physician I make the assumption that 45 percent antibiotics -- and that datum is valid, and if anything it is under-recorded -- 45 percent of patients with appendectomies is just too many. There are not that many complications in patients that need antibiotics in a routine appendectomy. Across the country we have a problem with the use of antibiotics.

I do not imply from that, and that I think is the misunderstanding, that therefore physicians' motivations are bad. I am saying the practice is not as good as it should be. That does not mean the motivations are bad. It just means that we have a problem of how to help them use their motivation. I think physicians are motivated. I have always felt that most are. The question is: how do we get them to be able to use what knowledge we can impart, and to practice what they admit is good medicine?

CHAIRMAN RISING: The literature has been full of the business of "antibiotic umbrella," prophylactic antibiotics are bad. This is not new. It has been going on now for years. We have had programs on it; we have had teaching on it; we have had things in which doctors in a pre-test and post-test would show a change of knowledge if they are answering what you told them. The point is they do not entirely believe you. It is pretty hard to take a doctor who was practicing before penicillin and convince him that penicillin is bad. It is just very hard. I happen to have been one of them. I practiced a good part of my life before we had antibiotics, and it took several proofs before I was willing to accept that, and a lot of doctors apparently still do not.

MR. JENKINS: Then what does that say to you as an educator? You are not being effective because you have not changed the behavior.

CHAIRMAN RISING: That is right.

MR. JENKINS: Then you failed as an educator.

CHAIRMAN RISING: Amen.

MR. JENKINS: That is what we are after here, I think.

CHAIRMAN RISING: Why don't they believe us?

MR. JENKINS: Have you gone back to find out? Did you go ask those men why they do not? Don't they trust you? You have been around in this area a long time, I would assume from what I hear. They are paying your fees.

CHAIRMAN RISING: We have actually talked with them and asked why they do not believe us, and they just say: "Well, you know, for five years I have not had any trouble with penicillin, antibiotics, etc. I have not had a case of pseudomembranous enterocolitis." Until they have one, they will not believe us.

MR JENKINS: What does that say to you? Does that not say you need another technique.

CHAIRMAN RISING: It sure does.

DR. BANK: I would presume that those who have partaken of the riotous activities on many campuses at some time or another had social studies or civics in high schools. Are you then saying that the high school situation is no good?

MR. JENKINS: I am saying that it has failed, absolutely, one hundred percent. I'll buy that!

CHAIRMAN RISING: I do not think anyone is going to argue that we do not fail and that we should not try to evaluate.

MR. JENKINS: Then we have to change it. That is what the public school has not done. It has not changed.

DR. CLEMENT BROWN: I want to respond to the antibiotic thing because we did a study on the use of antibiotics in our hospital. I am going to mention this and discuss it a bit because I think it gives you a little bit of an idea of how much I think needs to be done to really change physicians' behavior.

Because of some other studies we are doing it seemed to us that a lot of the antibiotics were being used inappropriately. We reviewed fifty consecutive uses of antibiotics by members of our medical department and we found that only 30 percent of the antibiotics used, in other words, fifteen instances with antibiotics indicated, in the proper drug, dose, route, duration used, by our criteria. Those were my criteria and

those of a general practitioner working with me, and then we brought someone in from the medical schools to react to what we thought. We told the people in the medical department what we found. We asked them each individually to take a slip of paper and write down No. 1 and No. 2. Next to No. 1 put down what it is they thought we found; and next to No. 2 put down where it is each thought he wanted to be in terms of appropriate use of antibiotics.

They did this individually. We collected them all. The average of where they thought they were was a 35 percent appropriate usage; and where they said they wanted to be was a 75 percent appropriate usage. We then revealed that they were at the 30 percent appropriate uses and we gave them some examples of what we felt were inappropriate uses, and they were willing to accept them.

That was a start, but still we did not know what we were dealing with, except a lot of inappropriate uses of antibiotics that they agreed on, and a great cognitive distance between where they were and where they wanted to be.

We then gave each member of the medical department an examination, which half of them flunked and half passed. The average score was 69 percent, which was less than the 75 percent they agreed to on a practice level, but it was more than twice as good as their actual practice level at 30 percent. In giving this examination, and forty-two of the forty-four members of the department took it, we made a further diagnosis of what some of their problems were -- lack of knowledge, lack of problem-solving skills, because we constructed a problem-solving type of examination.

We then held a series of conferences on antibiotics, using the examination primarily as the point of discussion. Now, just with respect to sore throats or management of pharyngitis, or whatever a doctor might call it when you give someone an antibiotic for a sore throat. We spent an hour-and-a-half in discussion on two different occasions with members of the medical department. I will not go into details of the kinds of things we discussed and the kinds of agreements to which we arrived.

To give one more example, sure, the physicians agree that we ought to get a culture on everyone's throat before we give him an antibiotic. You haven't any idea in the world by looking, and loads of studies have been done on this. You have to have a culture. It is the only way to find out. They have said in these discussions: "Yes, that is fine if you are at the hospital, but we are in our offices and we do not have our culture media, we do not have this."

In just this one area we developed a program whereby the lab would pick up from their offices every day any cultures. We began using these little culture tubes, which they could actually put in their pocket -- it will stay for twenty-four hours -- and they can drop it by the hospital the next day. We reduced the cost because we have anticipated, and we have come up with a three times greater use of throat cultures in our hospital from the out-patient service. We promised them we would reduce the cost for this, so that would not be much of a factor. We worked out some other kinds of arrangements. That is one little tiny area.

We have re-studied the situation, the use of antibiotics, and over-all we now find a 60 percent appropriate usage of antibiotics. It took a lot of activity, but they are still not where they want to be. In fact, before we told them that we asked them again: "Where do you want to be?" They felt at this point they really thought they ought to be at the 80 percent level. Still, we have a substantial distance to go. What we are now agreeing to is individualized small-group sessions, depending on the kind of problems we have determined from our evaluation of patient care, so that a group of physicians who are having problems, particularly with use of an antibiotic, will get a series of conferences or whatever it is that the learning experience happens to be.

Unless we get involved in doing this kind of thing to this extent, I do not think we can expect to see the kinds of changes that are really needed. You know, to have

had a one-hour session or even three hours on the use of antibiotics on the basis of the 30 percent appropriate usage that we originally started with, I do not think it would have had much effect.

Unfortunately, we did not do a controlled series and find out what it would have done if we had made the same kind of changes, but there are a lot of indications. For example, not a single dose of chloramphenicol was inappropriately ordered in six months after this conference, not a single use of the most popularly used antibiotic before the conference, it shall be nameless, which is an antibiotic that the disease that it should be used to treat has not been invented yet. Not a single dose of that, so we sort of figure that the series of conferences and discussions that we have had might have had something to do with this change. I will grant that we really do not know. It could have been just happenstance, you know; it might have happened anyway if we had never done anything.

In terms of the cost of this total program -- my salary, the girl who works with me, the cost of PAS, paying physicians $50.00 per hour for auditing charts -- all these factors are a cost to our hospital, and we have gone from primary appendectomies -- just to give you a couple of examples of changes we have achieved -- primary appendectomies going from a 51 percent rate of acute appendicitis to 81 percent rate acute appendicitis on primary appendectomies, which has held up over two years; going from 25 percent to a 13 percent complication rate in hysterectomies; and 30 percent to 60 percent appropriate use of antibiotics already mentioned; from one subtotal hysterectomy a month to one subtotal hysterectomy a year. The total cost of this program is $45,000 a year at our hospital, and less. With a hospital with a $6,000,000 budget, somehow it seems to me that this is not an inordinate cost for evaluation, the whole process, evaluation, education, re-cycled back to evaluation. I do not think it is that expensive. You can achieve change, but it takes what a number of people have been saying, particularly Keith Jenkins, it takes a lot more work on our part as educators, a little ingenuity, a little inventiveness, and mostly just a lot of hard work.

CHAIRMAN RISING: Dr. Brown, do you think this is an indication that this kind of precision of behavioral analysis and intensity of education and degree of cost-effectiveness is pretty uniquely appropriate in the community hospital as contrasted with a state institution that is.....

DR. CLEMENT BROWN: I don't know, Jesse, I don't know.

DR. BISH: Then the traditional organization of continuing education programs that is really in format a seminar, a walk-on-walk-off type of thing probably does not lend itself to the evaluation that you are suggesting? Is that right? This kind of an evaluation is not suitable in that context? Would that be your judgment?

DR. CLEMENT BROWN: I don't know. I don't know. I don't work in that context. I don't work like Jesse does here.

DR. BISH: What I am getting at, such a program certainly does not have that continued relationship between those who carry on the program and follow-up, the interaction, and some personal kind of dynamism is brought to relate to the age of modicification. It surely does not have that. Then, the conclusion is obvious, that we can't do that kind of an evaluation in our context?

CHAIRMAN RISING: Not that cheaply.

DR. CLEMENT BROWN: Probably not that cheaply, I don't know. I keep thinking there are some things that really make a difference that we could evaluate, even as people come from around Kansas to this center. There must be some things we expect to learn when they come here and they go back and refer patients here, and these patients still have the same old diagnoses, they are not using any of the new diagnoses. They have not gotten any of the new drugs that Dr. Rising has told them about. There must be some kind of indicator that you are having a substantial impact, or else that you are not having any impact.

DR. THOMAS BROWN: First, being the Brown here who is not a physician, and you have asked for nonphysicians to participate, I feel compelled to say something whether I have something to say or not.

CHAIRMAN RISING: A lot of us do that.

DR. THOMAS BROWN: I think we have deteriorated to some extent here because we are getting off the prime topic, unless we are all defining what the prime topic is. We need to get a delineation of your model that works in a community hospital under this circumstance, and we have to take a look at a delineation of something that works in a medical school in terms of evaluation, or something they are getting at Vanderbilt in terms of evaluation. I believe that we have to delineate these, we have to communicate these, we have to analyze these and look for the appropriate usage, and we have to generalize from this.

I want to make a plea for what Dr. Hugh Petersen is saying; I want to make a plea for what some of the others are saying, that we need to come together with these processes in some manner. We need to make it visible so that it can be replicated, so that it can be tested to see what works under what kind of circumstances. I would like to simply get my opinion in, that I think this is what we really need to address ourselves to.

DR. IRWIN BROWN: I wanted to ask Clem Brown one question before we go on, and I am sure that you followed this through, but it seems so simple to say that the incidence of appendicitis went from 51 percent to 80 percent. It would also be interesting to know what percentage of ruptured appendices there were when you had 51 percent and 80 percent. I do not know what it was, but I can give you an example: in the New England Journal of Medicine about 1950 there was an article on how one should make an absolute diagnosis of appendicitis in a child under the age of five. They had 100 percent accuracy on diagnoses, but they had four deaths in their small series.

In the hospital I was in I became interested in the subject, so I combed back through that age group, and there was about a 51 percent accuracy of diagnosis and zero mortality. There are so many things that are complex that I would not want the others here to think that a simple test of one thing, such as: "Do you do a throat culture?" Yes, I did a throat culture and I waited three days; therefore I got three days late in starting my treatment. I did a throat culture and I started them on antibiotics with a 75 percent or 90 percent chance of being right on antibiotics. What happened to the culture after they got it? What was the meaning of it? Those are all very complex things. You cannot simplify it.

CHAIRMAN RISING: I recently heard a microbiologist, an infectious disease person -- I won't say his name because I can't be sure -- discussing this. He hazarded the guess that the "inappropriate" use of penicillin was responsible for the remarkable decline in the incidence of acute rheumatic fever that has been seen recently. This was something that he said could not be proved or disproved, but it was an intriguing idea. Perhaps our best criteria are faulty! They often have been in the past.

DR. CLEMENT BROWN: I would like to respond to his question. In primary appendectomies it went from 52 percent to 81 percent and stayed at 81 percent the last two years. We had fewer complications and we have had no deaths for the five years I have been there from an appendectomy. We did not expect any.

We did thirty fewer appendectomies, which may be a little indication of something. Almost all of our surgeons do almost all of their surgery in our hospital, so we know that they are not taking their appendectomies elsewhere. We are now doing thirty fewer, we have fewer complications, but we know very precisely why the situation has changed, because of our referral system. The greatest portion of the problem was that our general practitioners and our interns were referring patients to surgery for an appendectomy, and they were being operated on in less than six hours of hospital admission. I think it is relatively rare that someone needs to have an appendectomy within six hours. Most

of the surgeons that I have a great regard for say that you can watch someone and very often things will change in that period of time and you will not need to operate, but if you were operating on them quick the signs would not change.

DR. EISELE: I want to compliment Clem Brown on the very sophisticated studies they have been doing and the effectiveness of them, but this is not entirely new. More than twenty years ago I was involved in a little seventy-five bed hospital which did a medical audit for the staff, and they immediately set up their own medical standards committee, twelve general practitioners on the staff at this hospital, the only hospital in the county, and they started reviewing our own work in a very unsophisticated manner. Three years later they invited me back. I found that the number of appendectomies was half of what they had been doing over a five-year period before the study -- before their committee started working. The number of tonsillectomies had fallen to one-half. Nonspecific pelvic surgery had dropped remarkably.

In very unsophisticated simple ways they are reviewing their own work and, I think, have effected very remarkable improvement in the quality of their work. They can be justly proud of this.

I had to put this in as my testimony that a medical staff in any hospital can organize itself to do its own evaluation of the quality of care and thereby educate the members.

DR. KORST: I would like to change the subject for a little while -- not that this is not very interesting. It is very refreshing for me to attend a meeting like this with a number of educators, administrative people, a good mixture, but as a physician I would like to make a point here, that much of our discussion over the last two days has been directed at what physicians can do in evaluation of programs. One of the greatest problems facing the physician today is the overextension of his time, and I am sure you are all aware if you read the Saturday Review, every article you read is the extension of the problems in manpower. Increasing students and physicians isn't the answer to this. The answer to this is the answer to the question: "How are all of the other supports, the allied supports in medicine, going to be helpful in the future?" I do not think we have talked very much about how you, the non-M.D. people in this conference, figure into this. Are we really just sitting around a table telling the physician what more to do in an already extended situation? Physicians need help, the whole system needs help, and I would like to get some answers to that.

MR. JENKINS: I will respond as someone who has not been around long enough to know what I should not be doing.

Because he is not going to accept our judgment, just as we teachers do not accept others telling us how to do our job. I want to work with you, as a physician, but I want to sit down and I want to throw out my ideas, I want them chewed up and we come out with a plan of operation. I can come in and do certain things, once they are acceptable to you, but if I ask you to come in and do an audit of your hospital, let alone your practice, you are immediately suspicious of me until you know me as an individual -- until you know that I can be trusted not to hand that information out to any place that it might be useful, to my own benefit. Therefore, the only thing that we can do, as I see it, is to build trust. That means getting involvement.

That means, if we are talking about a particular type of physician, a particular type of practice, or perhaps a group of academies, or one particular academy, get together and they work on it. For example: in my case in RMP I am going to try to talk to physicians, and if they trust me and trust RMP -- which I do not think they do right now -- because I represent RMP, they will eventually trust me and we can sit down and do something together.

I am very excited because I am going to be observing a program that Dr. Thomas Brown has been working on in Illinois next month, and I am supposed to go back then and do it under Dr. Miller's group as a contract under RMP. It is going to be a challenge

to me to be teaching physicians and hospital administrators how to do an audit of their hospital so they can plan their continuing education program. This is what I think we should be working toward, and I think it is what we have been saying. I think the whole key to this is the topic on physician attitudes, because if we do not have the physician we have nothing! It is dead, because without you gentlemen -- no matter how much we may talk about you -- without you we cannot do anything, we can never get anything done.

MR. COOLE: I would like to comment as to what the educator can do for the physician in his practice. I am an educator rather than a physician. First, there are certain ways that we can help you improve your efficiency of practice, if given the opportunity of finding out what happens now in your practice. I think there are techniques we can apply in terms of more efficient management, in terms of better record keeping. With respect to record keeping, I would like to add the idea that before we jump on the bandwagon of medical audit, let us realize that we may be evaluating the physician's record keeping rather than his patient care. This is the danger that we often overlook in the utilization of a medical audit.

Beyond that, with the Academy I have several proposals under way at this time, one of them being a manpower utilization in pediatrics to include not only the utilization of physicians but the utilization of nurses and other allied health workers. We need to consider the input from patients in this, and in pediatrics we certainly need to consider the parents. These are the kinds of things that educators can do to help physicians improve the efficiency and effectiveness of their practice. Before we can do this we have to be allowed to gather some baseline data. This is the plea that we make to you as educators: give us some baseline data to work with, and let us show you what might be possible, with your help.

DR. EISELE: Jesse, for twenty years I have heard this criticism that we are evaluating the physician's record keeping rather than his practice. Let me say, I cannot recall ever seeing an excellent record associated with bad practice, or vice versa. They generally go together pretty well. Poor records are associated with poor practice, and the excellent records are associated with excellent practice. There is a very good correlation. I cannot give you a quotient, but this is my observation.

I have observed or detected some slight evidence of hostility. In fact, I think there have been a few outright doctor haters in the room. Let me tell you right now, if you are an educator or an evaluator and have this kind of an emotion, you better get out of the business. You are licked.

MR. JENKINS: I don't think we are doctor haters.

DR EISELE: I wasn't looking at you.

MR. JENKINS: I know, but in the whole group, and I have talked with the others, and I do not think we are. We would not be here if we were, I hope not, and I really do not believe so. It is a matter of if you hit head-on and you hit head-on, after a while you become frustrated. What do you start doing? You start bitching. Now, if we can stop hitting head-on, open the door and talk to each other, we may find out that both of us are pretty nice guys.

DR. EISELE: My point is that the communications are just about zero when this type of attitude is....

MR. JENKINS: All right, let us both make an attempt to open the door and not both be traveling the other way.

DR. LESTER: I am a practicing physician and I came here to find out about this. By the way, I was told that everyone else was paid to come here, I was not.

I want to talk about this very thing. It is a concern to me, and was a concern before I came here, because I do think we should do a better job of education. I think someone like myself, or someone like Dr. William Fifer, who is a well-known practitioner who has now joined the RMP, can open doors that others cannot.

I just want to say that these veiled threats of third-party intervention and all that are fruitless. I seriously doubt that they can do the job if we can't. They're not going to do it better because without our cooperation, as all of you have said, it cannot be done. I want you to know the whole effect of these little innuendoes on your future cooperation. This should not be part of your vocabulary. You should wipe the expression off your face. You should act as though you have heard of it, but give it no credence, because we also read Saturday Review and all these other things. There is plenty that doctors can say about the motivation of other people -- electricians, carpenters, taxi drivers, and everybody else that provide services. We all have feelings about senators and their efficiency, and about government and its effect on inflation. The point being, however, and I hear fellow physicians say this: "What we really need to do by education is have some sort of a club." Academies have this and we tie membership to continuing effort. When I hear in the background, "you've got to do all this for these notable reasons, and if you do not, one day we are going to take your license," it has the same effect as the eighteen-year-old girl who comes in my office pregnant, who is not married, and who checks me out with two or three visits to see whether she can trust me with information, or with herself, before she gets around to discussing her pregnancy. They want to make sure that in the next room they cannot be heard when they discuss their private problems with me. If they were perfect to start with, they would not have to see me as a physician. They do not need me. People who are imperfect need to come into my office, and their trust with what they tell me will be held in confidence is all important.

From a practical standpoint, I have been waiting to hear, and it has not been mentioned: how do you assure, Dr. Brown, that the information you gather cannot be used by someone else?

DR. CLEMENT BROWN: I do not know. No one seems terribly concerned, in our staff at least.

CHAIRMAN RISING: For one reason, you have not published it. Most educators publish their evaluations if they have been successful, at least.

DR. CLEMENT BROWN: We do have an article coming out in JAMA.

CHAIRMAN RISING: There may be a problem after it comes out.

DR. CLEMENT BROWN: No, I don't think so, because I have had board trustee members and our staff members at Dr. Eisele's meeting and we have talked about it to some extent.

Let me respond to the business of trust, which does seem to be a matter of concern right at the moment. On this little opinion survey that was taken yesterday: "Most practicing physicians have little desire to keep up with the advances in medical practice." Only two people here agreed with that statement; 45 disagreed with it. I think most people here sort of trust the motivations of a practicing physician, so I don't think that is really too much of a problem. When I first came I thought, "Oh boy! Here we go again; another two-day session of a bunch of educators and we have no clearly stated objective," and the like kind of things that educators are supposed to say, I guess. Jesse, you stated the objective, and did you set my teeth on edge! You said: "The purpose of this conference is to write a book for -- Health Manpower" -- or something like that, "and not to help those in attendance. We are all experts in what we are doing, and we are here to sort of write this report, book, or whatever." Like Wow! I don't know why you get me so mad so quick at times, Jesse.

CHAIRMAN RISING: I don't know why you are so irritable, Clem.

DR. CLEMENT BROWN: Then, you know, we have been through this two days now, and I think back and I say, "what has happened?" We had some group discussions yesterday afternoon and the session this morning. Now, what we seem to be doing is diagnosing our needs as evaluators of continuing medical education. What are our problems, and how are we going to get around them? We have listed them under physician-learner problems, educator attitudes, costs, and measurement problems.

Right now I am wondering, "have we done enough diagnosis these last two days of what our problems are as potential CME evaluators so the therapy is obvious to us, because at least in our hospital we have often found that once you have performed the diagnosis the therapy is quite obvious.

Here is where the real question comes in: "how will we know if our diagnosis has been sufficient and the therapy is obvious to ourselves as continuing medical education evaluators, because we have just been through a two-day session as evaluators and we are not going to evaluate it at all?"

I would propose that we evaluate what we have done the last two days by having Jesse Rising, if he would, mail out this opinion survey that everyone took the first day. It would be interesting to see if there are some differences of opinion now. All of you have your original data. Even more important than that, if we have people not fill out these things on the bottom which really get at attitudes, but take No. 18: "I have developed a curriculum for at least one of the CME programs I have conducted by first gathering data substantiating patient care deficits," I would like to know if 20 say "yes" or 30 say "yes" six months from now. I would like to know a year from now how many will say: "I have gathered data representing before and after, and I have conducted a program based on definite and significant improvement in patient care." I would like to see if we have, in fact, moved to doing anything as a result of this two-day session.

Also, I would like to propose that we modify this and see if anything happens. Right now, fill out the top portion -- and mail it back in -- then, do a six-month or one-year follow-up because, we are talking about evaluation and we have not evaluated what we have done. We may need another session like this a year from now. We may have found out in that time that no one is doing any more evaluating than what we already have data on at this time, and we will have to realize that whatever happened today and yesterday did not help us to get ahead with the job, and we need another session like this, but better, or different -- I do not know.

CHAIRMAN RISING: Clem, why do you insist on changing my objective. I did not have an objective of educating this group. I told you that in the first place. This was not to educate you or to change your behavior, but to expose the ideas and thinking of those who are experts and practitioners of some phase of continuing education.

DR. CLEMENT BROWN: That is what you told them when I got here; I had some of my own objectives written out.

CHAIRMAN RISING: But those were <u>your</u> objectives.

DR. CLEMENT BROWN: No. I shared them with my group yesterday, and they turned them all down, but that is fine.

CHAIRMAN RISING: This was not planned as an educational exercise for us. I told you that. I made it perfectly clear in the letter I wrote. The objective was to get opinions of people who are "pros." Everyone here is presumably a professional in some phase of continuing education or evaluation. If that isn't true, you came under false colors, because the letter did not go to anyone who was not. We wanted to get a book of opinions that can be put in the hands of others. I have already had people here tell me they would like to use some of the things from this conference that will be useful to them if they can be quoted. So, don't change my objective, please, Clem.

DR. CLEMENT BROWN: Suppose no one here is doing any more evaluating a year from now than what they are doing now?

CHAIRMAN RISING: I really don't care. That's their business -- their choice. That has nothing to do with my objective. Some people may have been convinced for or even against by this...

DR. CLEMENT BROWN: Most people that are here are not even doing any evaluating, by their own admission. How can we serve as experts to the rest of the world?

CHAIRMAN RISING: I did not say they were all experts in evaluation. I said they were all experts in some phase of continuing education. By your definition, of course, they are not doing continuing education if they do not do your kind of evaluation, which just possibly is a little narrow-minded.

Clem, I agree 100 percent with everything you said relative to your circumstances and your situation. I agree wholeheartedly with you that the over-all objective of continuing education is to change behavior. But the objective of the Minnesota Academy of General Practice's program -- even though it is partly fellowship and organizational maintenance -- the objective of their educational part of the program is not necessarily an immediate change of behavior. It may be a third or fourth step back from that, by changing their members' approach, by keeping their interest in education, and by stimulating their interest. This is one of the problems of applying Hugh Petersen's criteria to absolutely everything until he can show us how we can measure change of attitude. Theoretically, it is possible, but the tools have not been here. The evaluations have not been evaluated.

Please, let us not be paralyzed because we cannot evaluate everything in Hugh's terms. In writing a proposal for RMP it is really important to write objectives that are easy to evaluate in terms of a change of behavior or attainment of the objectives, so that you have a proposal that will be supportable by a renewal. Is this right, Thelma?

MRS. SCHNEIDER: It would help, but it would also help to show that one day we might get around to patient care. Everything contributes toward that goal.

CHAIRMAN RISING: This, of course, is another point. If we are talking about evaluating patient care, are we too late? It has been pointed out that we cannot eliminate mortality. People are going to die. We can improve longevity very little unless we do away with degenerative diseases that no one, no matter how good an educator, knows how to do.

To improve patient care we are dealing with sick people because only sick people are patients. How do we prevent illness? Ideally, no one would be sick until the day he dies. This is the goal we are really aiming at in health care -- the prevention of illness, the prevention of any sickness.

If we have a pie-in-the-sky approach to what our objectives must be we are going to have people say: "Well, phooey, it's useless. I just can't do it."

DR. CLEMENT BROWN: Did you just say that we are not going to evaluate by finding out what people are doing six months or a year from now?

CHAIRMAN RISING: I don't really care. This particular group of people has already heard all these arguments before, you know. You are not going to change their behavior dramatically, because of hearing them once more. They probably are doing all they can. If they are not it is because they have not been convinced, just like those who have been hearing about antibiotics. Unless we have control of this group in some way so we can control their behavior as you can in your hospital, we are not going to make the Minnesota Academy of General Practice do something that it cannot see how in the name of common sense it can do.

DR. PERLMAN: Sure it can.

CHAIRMAN RISING: How?

DR. PERLMAN: He has availability of 25 percent of the patient discharges in Minnesota today. They have, not the details, but rough statistical data on 25 percent of the patients in Minnesota right now.

CHAIRMAN RISING: How is Dr. Lester going to say that the Academy of General Practice has made this change by virtue of their program?

DR. PERLMAN: I do not think that is important.

CHAIRMAN RISING: Well, this is what some appear to have been saying.

DR. PERLMAN: I said earlier that I would be satisfied -- if we are going to try to prove things absolutely scientifically, I don't think we are ever going to get off the ground with any of these programs -- I think we are going to have to be satisfied with relationships. We are going to have to be satisfied that L-dopa is a good drug if given in a higher, much greater amount and, hopefully, given properly a year from now. I don't think it matters, because this is such a fast-moving thing. I don't think it matters right now that we are going to have to try to give a percentage to what journal ads versus an article in a journal versus a course that someone ran, what percentage each was responsible for the increase.

CHAIRMAN RISING: That's correct, Jack, but, you see, Dr. Fifer is not going to be able to take credit for this because it would have happened anyway.

DR. PERLMAN: Why should he want credit? He says he does not want to get credit.

CHAIRMAN RISING: I know he doesn't want credit, but some of those here seem to think that the Minnesota Academy must have either credit or blame.

DR. LEMON: The purpose of evaluation is to find out which method is effective. The increase of L-dopa may be because of the increased activity of details and have nothing to do with Dr. Fifer's program.

DR. PERLMAN: That's right, and I think there are some programs we cannot evaluate to get this definite proof.

CHAIRMAN RISING: Would you please repeat that?

DR. PERLMAN: There are some of these programs that we are not going to be able to evaluate because they move too fast.

CHAIRMAN RISING: Is that any reason why we should not put them on? Are we "perpetrating a fraud?"

DR. PERLMAN: I think you can take the same kind of program on something that is not as fast-moving and as dramatic and see if it does make an effect. You can just apply it, you can just take a "guesstimate," an educated guess, of whether this is going to have an effect. Is this some drug that is so dramatic that it makes people so much better?

DR. CLEMENT BROWN: You always do evaluate to see what kind of an effect you have had in continuing education. I would like to evaluate the use of L-dopa at some point a year or so from now to see if 70 percent of the people are using it inappropriately. I would then construct an educational program to get at the need. If 90 percent of them, 80 percent, whatever figure we want to accept, are using it appropriately, let's forget it. Let's spend money on something else.

CHAIRMAN RISING: Everyone here is getting a bit restless. The unanimity that seemed to have developed in the discussion groups yesterday afternoon is evaporating under the stress of fatigue, and flight departures are getting closer and closer. I will ask for one final comment and I think we will then have to adjourn.

MR. SCHMALGEMEIER: This is the first time I have been able to have the last word in a long time -- my wife is at home.

CHAIRMAN RISING: It is now on the record.

MR. SCHMALGEMEIER: We keep coming back to one thing, and that is the fact that we have to establish the objectives for which we are carrying on our educational program.

The educational program, or activity, in which we are involved here in Kansas City has one objective, very easily evaluated; everyone here will be able to tick it off at the time it is completed when the book is published. I think, Clem, that this is in a sense the basic behavioral objective of this course. I wish it were otherwise, but that is a different situation. It can easily be checked when it is published.

I think when we are talking about the over-all pattern of the educational evaluation we are talking about, we have to tie it down to the kind of objectives we have. I wonder what Dr. Fifer's objective is in offering a course on L-dopa at that time. If it is to increase the usage of L-dopa he is going to be able to say: "Yes, the increase in usage in L-dopa came about." If he is going to say it is to get increased effective use of it, then he must do some evaluation of what the behavior was before and what the behavior is going to be afterwards. You cannot do it in a vacuum. If it is a head count you are looking for, forget the course, save the money, spend it on something else, have a good vacation, go to Bermuda -- I do not care what it is -- the use will go on just the same.

We are back to the same thing again that we talked about at the beginning of the program; that is, let us establish our basic objectives. Let us establish the kind of evaluation techniques which will result from our objectives and then let us test the objectives to see whether we have carried them out. I think that many of us, and I think that most of us, would say that those objectives would have to be in some kind of action rather than some passive result. Therefore, the behavioral type of output or objective is what we are looking for. We do want changes in processes and procedures. These things are available in math courses, such as we are carrying on in Michigan, and others are carrying on through the departments of postgraduate medicine. They are available through the hospital-oriented programs, such as Clem Brown is working on. They are also available through some of the other activities.

May I put out one suggestion: if we cannot measure it, let's not do it at all.

CHAIRMAN RISING: Someone in the audience just said: "I hate to close on that note -- if we cannot measure it, don't do it!" This makes the untenable assumption that, if a program isn't evaluated, it ipso facto is not educational.

The last word will be that I want to thank the consultant planning group, Gail Bank, Clem Brown, Irwin Brown, Hugh Petersen, and Lee Stauffer (who, unfortunately, could not be with us) for helping us plan the format for this conference. If it has not been successful, I will take the blame. If any of you have enjoyed it, or had any fun, which to me is always an objective though not stated, or if there is anything that comes out in the book that is useful to anyone, it is to their credit and to the credit of all of you. Thank you all for being here with us. It has been a real pleasure to host this conference -- it was not an educational exercise, it was a conference -- it has been a real pleasure to host it and to meet all of you. I have thoroughly enjoyed myself. I love a good argument. I apologize if I have talked too much, and I know I have. Thank you all again.

Appendix A

EVALUATING THE EVALUATORS

Ethel Nurge, Ph.D. (Anthropology)

It was not my privilege to attend the conference on evaluation in continuing medical education, inasmuch as I was away from the Medical Center at the time that it was held. On my return Dr. Jesse D. Rising told me about the conference and made available to me the program and the proceedings of the second day. I read both of these with a great deal of interest and on Dr. Rising's suggestion have written the following as a commentary after, but I hope relevant to, the symposium.

Postgraduate medical education is not new of course and yet the emergence of the field as a field is a recent phenomenon. Even more recent and marked is the acknowledgement of need for objective evaluation and a self-conscious and sophisticated search for evaluative techniques. In telling you this, I do not inform you. Perhaps what I can say, that will give some perspective and a new orientation, has to do with the nature of social change and man as a social animal. Man as an innovator both creates and resists change, and both of those aspects, man creating and resisting change, are evident in the papers and discussion of the conference on evaluation in continuing medical education.

When a novelty, innovation, or new behavior is urged on a group, their reaction to the suggested change may be as wide as is the composition of the group, but the chances are that the responses will not be so wide-ranging. In fact, it may be helpful to think in terms of a series of selected responses, selected reactions to suggested change, i.e., the inauguration of evaluation as a standard and continuing procedure. It may be further helpful to think in terms of human beings as falling into certain categories as they are judged acceptors or rejectors of change. A wider understanding of man as an innovator or, conversely, staunch defendant of the status quo may help us to understand, predict, and effect social change.

To discuss those who accept or reject an idea we will begin by considering two interlocking hypotheses about the nature of change. The first is that an individual will not accept a novelty unless he is convinced that it satisfies a want better than some existing means at his disposal (Barnett 1953:378). I take this to mean that the body of physicians as a whole will not welcome and support surveillance, evaluation, measurement of their behavior, performance, and practice unless they are clearly convinced that it does something for them that peer review or self-appraisal cannot do. I think that the participants of the present symposium have amply demonstrated that objective evaluation can valuably supplement peer review and self-appraisal. Indeed there are areas (which I leave to the experts to delineate) which only outside evaluation can effectively cover. But the first hypothesis is that an individual will not accept a change unless he is convinced that it satisfies a want better than some existing means at his disposal.

The second hypothesis is that there are biographical determinants antecedent to lack of satisfaction as a pervasive continuing state in a certain number of individuals. These individuals may have had specific and repeated disappointments which are a genesis for long-term pervasive dissatisfaction; it gives them a perspective which colors their views of large but variable sectors of their culture (Barnett 1953:379). It is these very individuals who, providing they are healthy, are most open to suggestions for change. In other words, there are individuals who have predisposition to change and those who do not, and one lesson for us may be that we need to think of so training our medical students that they are less satisfied and more malleable and more open to change. I think this touches on a peculiarly sensitive point in medical education since there is a traditional and, in many situations, a valid need for an assertive, authoritative posture and character on the part of a physician. Such a posture, in most instances, is successfully modeled, copied, and well learned in medical school. To teach, at the same time, a contrary kind of behavior as a result of which the individual will not rest unassumingly on his own interpretation, but rather constantly seek a reason and a way to change it, presents a dilemma. In other words, perhaps we have to find a way to make

the medical students dissatisfied with themselves at the same time they are being taught to act as serene, knowledgeable characters. This is neither easy to teach nor easy to practice. There is very little in the medical student's training or experience that makes him humble; there is a lot that makes him the contrary.

At this point, let me enter a caveat. There are no clear-cut categories of individuals who are acceptors as opposed to rejectors; no one person is wholly or persistantly an acceptor or rejector. He is simply more or less of one than the other. Still we may be able to make useful distinctions among the acceptors. Looking at those who are more commonly acceptors, we find four categories. They are the dissident, the indifferent, the disaffected, and the resentful.

(1) The Dissident. In every society there are individuals who have consistently refused to identify themselves with some of the norms and customs of their group. They may give lip service to avoid punishment but the more courageous and independent these dissenters are the more they openly rebel and withdraw from participation. Insofar as they withdraw from the expected participation we may look for alternate forms of behavior, that is, we may hope for acceptance of, and initiation of, change.

(2) The Indifferent. Many individuals are prepared to accept a new idea of behavior because they have not dedicated themselves irretrievably to a custom or an ideal. In such individuals the receptive attitude is not due to dislike of an existing convention. In fact, they may be participants in the behavior to be changed, but they are not very enthusiastic about the existing convention. They do not completely identify with the ideas and they do not get much satisfaction from them. These people are the indifferent and are a group whom we may expect to influence when seeking to change behavior. The greatest number of individuals in this category are young. Indeed the best example is children who are notably indifferent to the values and passions of their elders. For change targets among physicians, then, we should perhaps look to the youngest ones to find those who are indifferent supporters of the status quo.

(3) The Disaffected. Some people start out with a firm commitment to, and great satisfaction in, certain behaviors promulgated by their culture but then, somewhere along the line, they acquire a distaste for them. These are the disaffected and they too are a category of physicians whom we may seek out as being most vulnerable to, and accepting of, culture change. In propitious circumstances, presenting such individuals with alternative modes of behavior is enough to wean them away from their previous convictions. In our present circumstances that may very well mean individuals who have been first satisfied with medical care and then quite strongly critical of it are the persons most likely to be willing recipients of critics' charges for evaluation of continuing medical education and practice, and will, perhaps, take to heart Pogo's confession, "I have met the enemy and he is us."

(4) The Resentful. In every society there are those who are resentful and this is our fourth category of change target people. All societies have a limited number of the top prizes, however such prizes may be defined, and always there are some who cannot get these prizes. One of the segments of a range of possible responses is resentment. These individuals are not resigned to their fate and, by contrast with complacent individuals of whom they may be envious, they are markedly receptive to suggestions for change. Therefore, when seeking targets for change behavior, seek among those who are getting the least and wanting something better, maybe lower echelon physicians.

Man as a social animal is not unpredictable. Within limits, sometimes easily delimitable and sometimes not, we may speak of his propensities. In this paper I have taken as my problem a formulation of categories of potential acceptors of change. The specific change being desired is the acceptance of evaluation of medical education, not only by the medical educators -- for they are fast moving to this themselves -- but also by other personnel.

Barnett, H. G., Innovation, McGraw-Hill Series in Anthropology, New York, 1953.

END

DEPT.OF HEALTH EDUCATION AND WELFARE

U.S. OFFICE OF EDUCATION

ERIC

DATE FILMED

9 27 71